



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:40:02 ; Search time 91 Seconds

699.443 Million cell updates/sec

Title: US-09-996-008B-2

Sequence: 1 MAKVLCVLYDDPVDGYPTY.....YSKGNATGGSEEAKEKKAV 401

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

```
Minimum DB seq length: 0
```

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 08

Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
24	/SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	2135	100.0		401	23	AAU991122	Mycobacterium vacc
2	2125	99.5		401	23	AAU991112	Mycobacterium vacc
3	2125	99.5		401	23	AAU991113	Mycobacterium vacc
4	2125	99.5		401	23	AAU991115	Mycobacterium vacc
5	2116	99.1		401	23	AAU991108	Mycobacterium vacc
6	2116	99.1		401	23	AAU99110	Mycobacterium vacc
7	2116	99.1		401	23	AAU99117	Mycobacterium vacc
8	2115	99.1		401	23	AAU99109	Mycobacterium vacc
9	2115	99.1		401	23	AAU99111	Mycobacterium vacc

10	115	99.1	401	23	AAU99114
11	215	99.1	401	23	AAU99116
12	2107	98.7	401	23	AAU99119
13	2106	98.6	401	23	AAU99118
14	2105	98.6	401	23	AAU99107
15	2105	98.6	401	23	AAU99120
16	2105	98.6	401	23	AAU99121
17	1965	92.0	401	23	ABW75656
18	1962	91.9	401	19	AAU781857
19	1710	80.1	399	21	AAU86650
20	1051	49.2	192	23	AAU49475
21	872	40.8	384	23	ABB93475
22	838.5	39.3	376	22	AAU60245
23	823	38.5	365	22	AAU65273
24	823	38.5	365	22	AAU63574
25	823	38.5	362	12	AAU94761
26	818	38.4	364	18	AAU16416
27	781	36.6	343	22	AAU37388
28	781	36.6	343	22	AAU37557
29	780.5	36.6	335	22	AAU34330
30	547	25.6	319	21	AAU04710
31	547	25.6	321	21	AAU04709
32	547	25.6	340	21	AAU04708
33	357.5	16.7	396	22	AAU96379
34	345.5	16.2	337	22	AAU96822
35	337.5	15.8	305	22	AAU96215
36	312	14.6	432	22	AAU65151
37	310.5	14.5	530	22	AAU91161
38	309.5	14.5	345	20	AAU31651
39	309.5	14.5	345	20	AAU22644
40	309.5	14.5	530	20	AAU31649
41	309.5	14.5	530	20	AAU22646
42	309	14.5	539	23	ABP40273
43	307	14.4	530	20	AAU31650
44	307	14.4	530	20	AAU22647
45	306	14.3	531	24	ABU19179

ALIGNMENTS

XX	AAU99122 standard; Protein; 401 AA.
XX	AAU99122:
XX	24-SEP-2002 (first entry)
XX	Mycobacterium vaccae formate dehydrogenase protein.
XX	Formate dehydrogenase; alcohol production; ketone;
XX	mutant.
XX	Mycobacterium vaccae.
XX	EP1211316-A1.
XX	05-JUN-2002.
XX	27-NOV-2001; 2001EP-0128170.
XX	29-NOV-2000; 2000JP-0363894.
XX	24-AUG-2001; 2001JP-0254631.
XX	(DAIL) DAICEL CHEM IND LTD.
XX	Mitsubishi K, Yamamoto H, Kimoto N;
XX	WPI; 2002-464925/50.
XX	N-PSDB; ABR86863.
XX	New mutant forms of formate dehydrogenase derived from Mycobacterium

PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX
 PS Example 1; Page 23-24; 42pp; English.

CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents the *Mycobacterium vaccae*
 CC formate dehydrogenase protein of the invention. This sequence was used
 CC to generate the mutant proteins of the invention.

XX Sequence 401 AA;

Query Match 100.0%; Score 2135; DB 23; Length 401;
 Best Local Similarity 100.0%; Pred. No. 1.5e-212;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPDYGYKTKTARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSSEGL 60
 Db 1 MAKVLCVLYDDPDYGYKTKTARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSSEGL 60
 QY RPYLESNGHTLVVTSKDGSDVSFERELVDADVVISQPFMPAYLTPEIRIAKANLKLALT 120
 Db RPYLESNGHTLVVTSKDGSDVSFERELVDADVVISQPFMPAYLTPEIRIAKANLKLALT 120
 QY 121 AGGSDHVDQSAIDRNVTVAEVTYCNISVAEHVMMILSVRNYLPSEHMAKGGWNT 180
 Db 121 AGGSDHVDQSAIDRNVTVAEVTYCNISVAEHVMMILSVRNYLPSEHMAKGGWNT 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDHRRLPESYEKELNLTWHA 240
 Db 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDHRRLPESYEKELNLTWHA 240
 QY 241 TREDMPVCDVVTNLNCPLEPTEHMINDETLKFKRGAYIVNTARGKLCRDVARALES 300
 Db 241 TREDMPVCDVVTNLNCPLEPTEHMINDETLKFKRGAYIVNTARGKLCRDVARALES 300
 QY 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 Db 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHSYSKGNATGSEEAFFKAV 401
 Db 361 PIRDEYLIYOGGALAGTGAHSYSKGNATGSEEAFFKAV 401

RESULT 2
 AAU99112
 ID AAU99112 standard; Protein; 401 AA.

XX AAU99112;

DT 24-SEP-2002 (first entry)

DE *Mycobacterium vaccae* formate dehydrogenase C146S mutant.

XX Formate dehydrogenase; alcohol production; ketone;
 KW mutant; mutleln.

XX *Mycobacterium vaccae*.
 OS Synthetic.

XX Key Location/Qualifiers

FT MISC-difference 146

XX /note= "Wild type Cys substituted by Ser"

PN EP1211316-A1.

XX

PD 05-JUN-2002.
 XX 27-NOV-2001; 2001EP-0128170.

PR 29-NOV-2000; 2000JP-0363894.

PR 24-AUG-2001; 2001JP-0254631.

PA (DAI) DAICEL CHEM IND LTD.

PI Mitsuhashi K, Yamamoto H, kimoto N;

DR WPI; 2002-464925/50.

PT New mutant forms of formate dehydrogenase derived from *Mycobacterium*
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials

PS Claim 8; Page -; 42pp; English.

CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.

XX Sequence 401 AA;

Query Match 99.5%; Score 2125; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1.6e-211;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPDYGYKTKTARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSSEGL 60
 Db 1 MAKVLCVLYDDPDYGYKTKTARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSSEGL 60
 QY RPYLESNGHTLVVTSKDGSDVSFERELVDADVVISQPFMPAYLTPEIRIAKANLKLALT 120
 Db RPYLESNGHTLVVTSKDGSDVSFERELVDADVVISQPFMPAYLTPEIRIAKANLKLALT 120
 QY 121 AGGSDHVDQSAIDRNVTVAEVTYCNISVAEHVMMILSVRNYLPSEHMAKGGWNT 180
 Db 121 AGGSDHVDQSAIDRNVTVAEVTYCNISVAEHVMMILSVRNYLPSEHMAKGGWNT 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDHRRLPESYEKELNLTWHA 240
 Db 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDHRRLPESYEKELNLTWHA 240
 QY 241 TREDMPVCDVVTNLNCPLEPTEHMINDETLKFKRGAYIVNTARGKLCRDVARALES 300
 Db 241 TREDMPVCDVVTNLNCPLEPTEHMINDETLKFKRGAYIVNTARGKLCRDVARALES 300
 QY 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 Db 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHSYSKGNATGSEEAFFKAV 401
 Db 361 PIRDEYLIYOGGALAGTGAHSYSKGNATGSEEAFFKAV 401

RESULT 3
 AAU99113
 ID AAU99113 standard; Protein; 401 AA.

XX AAU99113;

DT 24-SEP-2002 (first entry)

DE *Mycobacterium vaccae* formate dehydrogenase C256S mutant.

XX

XX	Formate dehydrogenase; alcohol production; ketone;
KW	mutant; mutEtn.
OS	Mycobacterium vaccae.
OS	Synthetic.
FM	Key Location/Qualifiers
FT	Misc-difference 256 /note- "Wild type Cys substituted by Ser"
XX	EPI211316-A1.
PD	05-JUN-2002.
XX	27-NOV-2001; 2001EP-0128170.
XX	29-NOV-2000; 2000JP-0363894.
PR	24-AUG-2001; 2001JP-0254631.
XX	(DAIL) DAICEL CHEM IND LTD.
PA	Mitsubishi K, Yamamoto H, Kimoto N;
XX	WPI: 2002-464925/50.
DR	New mutant forms of formate dehydrogenase derived from Mycobacterium
PT	vacciae have enhanced activity in the presence of organic solvents and
PT	are useful for producing alcohols from ketone raw materials -
PS	Claim 8; Page -: 42pp; English.
CC	This invention relates to a mutant form of the Mycobacterium vaccae
CC	formate dehydrogenase protein which has strong formate dehydrogenase
CC	activity in the presence of an organic solvent. The mutant formate
CC	dehydrogenase protein of the invention may be used to produce alcohols
CC	using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC	proteins, the mutant polypeptides of the invention retain high enzymatic
CC	activity in the presence of the alcohol product and so provides a higher
CC	product yield. The present sequence represents a Mycobacterium vaccae
CC	formate dehydrogenase protein mutant of the invention.
XX	
XX	
SQ	Sequence 401 AA:
	Query Match 99.5%, Score 2125; DB 23; Length 401;
	Best Local Similarity 99.8%; Pred. No. 1.6e-211;
	Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 MAKVLCLVLYDDPVNGYKRTYARDDLPKTDHPGQILPTPRAIDTPCOLLGSVSGELGL 60
DB	1 MAKVLCLVLYDDPVNGYKRTYARDDLPKTDHPGQILPTPRAIDTPCOLLGSVSGELGL 60
QY	61 RPYLESNGHTLVMTSDKDKDPSPVERELVDADVYSOPFWPAYLTLPRIAKKNKLALT 120
DB	61 RPYLESNGHTLVMTSDKDKDPSPVERELVDADVYSOPFWPAYLTLPRIAKKNKLALT 120
QY	121 AGIGSDHYDLQSAIDRNVTAEAVTYCNSISVAEHVMMILSLVRNYLPSHEMARKGNMI 180
DB	121 AGIGSDHYDLQSAIDRNVTAEAVTYCNSISVAEHVMMILSLVRNYLPSHEMARKGNMI 180
QY	181 ADCVSHAYDELAHMGVYAAGRIGLAVLRRLAPEDVHLHYTDRHRLPESVEKEMLTWHA 240
DB	181 ADCVSHAYDELAHMGVYAAGRIGLAVLRRLAPEDVHLHYTDRHRLPESVEKEMLTWHA 240
QY	241 TREEMYPVCADVNTNCPLEPTEHMINDETLKLFRRGAYTYNTARGKLCDDDAVARALES 300
DB	241 TREEMYPVCADVNTNCPLEPTEHMINDETLKLFRRGAYTYNTARGKLCDDDAVARALES 300
QY	301 GRLAGYAGDWFPQAPARKDHPWRTPMYNGMTPHISGTLTLAQARYAAGTREILLECFFEGR 360
DB	301 GRLAGYAGDWFPQAPARKDHPWRTPMYNGMTPHISGTLTLAQARYAAGTREILLECFFEGR 360
QY	361 PIRDEYLIVGGALGTGAHSYSKGNATGCSEBAKKFAV 401

Query Match	Best Local Similarity	99.5%: Score 2125; DB 23; Length 401;
Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
1 MAKVLVLYDDPVGVCYPTKTYARDDLPKIDHYPGGIILPTPKAIDFTPGQLGSVSGEGL 60		
1 MAKVLVLYDDPVGVCYPTKTYARDDLPKIDHYPGGIILPTPKAIDFTPGQLGSVSGEGL 60		
61 RPYLESNGHTLVTSKDKDPDSVERELVADAVVISQFPWPAYLTPERIAAKNLKLTAL 120		
61 RPYLESNGHTLVTSKDKDPDSVERELVADAVVISQFPWPAYLTPERIAAKNLKLTAL 120		
121 AGTGSDDVADLSAIDRNVTVAEVYCNISVAEHVYMMTISLVNRYLPSHEMARGGNNI 180		
121 AGTGSDDVADLSAIDRNVTVAEVYCNISVAEHVYMMTISLVNRYLPSHEMARGGNNI 180		

QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYTDHRLRPESVEKELNLTWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYTDHRLRPESVEKELNLTWHA 240
 QY 241 TREDMPVCDVYVTLNCPHPETEHMINDETLKFKRGATVYNTARGLCDRDVAVALLES 300
 DB 241 TREDMPVCDVYVTLNCPHPETEHMINDETLKFKRGATVYNTARGLCDRDVAVALLES 300
 QY 301 GRLAGYAGDVWFPPQAPKDPHMTMPYNGMTPHISGTTTLTAQARYAAGTREILCEFEGR 360
 DB 301 GRLAGYAGDVWFPPQAPKDPHMTMPYNGMTPHISGTTTLTAQARYAAGTREILCEFEGR 360
 QY 361 PIRDEYLIYOGGALAGTAGHASYSGNATGSGEAAKFKKAV 401
 DB 361 PIRDEYLIYOGGALAGTAGHASYSGNATGSGEAAKFKKAV 401

RESULT 5
 AAU99108 ID AAU99108 standard; Protein; 401 AA.
 AC AAU99108;
 XX
 DT 24-SEP-2002 (first entry)
 DE Mycobacterium vaccae formate dehydrogenase C6A/C256s mutant.
 KM Formate dehydrogenase; alcohol production; ketone;
 XX mutant; mutein.
 OS Mycobacterium vaccae.
 OS Synthetic.
 XX

Key Location/Qualifiers
 FT Misc-difference 6 /note= "Wild type Cys substituted by Ala"
 FT Misc-difference 256 /note= "Wild type Cys substituted by Ser"
 FT
 XX
 PN EP1211316-AL.
 PD 05-JUN-2002.
 XX
 PE 27-NOV-2001; 2001EP-0128170.
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAI) DAICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from Mycobacterium
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX
 PS Claim 8; Page -: 42pp; English.
 XX

This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.
 CC
 SQ Sequence 401 AA;

Query Match 99.1%; Score 2116; DB 23; Length 401;

Best Local Similarity 99.5%; Pred. No. 1.4e-210;
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAKVLGVLDPPDVGPKTYARDDLPKIDHYPGQILPPIKAIIDFPGLGSSVSGELI 60
 DB 1 MAKVLGVLDPPDVGPKTYARDDLPKIDHYPGQILPPIKAIIDFPGLGSSVSGELI 60
 QY 61 RPYLESNGHILVYTSKDGSDPSFERELVDADVVISQPPWPAVLTPPERIAKAKNLTALT 120
 DB 61 RPYLESNGHILVYTSKDGSDPSFERELVDADVVISQPPWPAVLTPPERIAKAKNLTALT 120
 QY 121 AGISDHPVDLSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNLPSEHARKGAWN 180
 DB 121 AGISDHPVDLSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNLPSEHARKGAWN 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYTDHRLRPESVEKELNLTWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYTDHRLRPESVEKELNLTWHA 240
 QY 241 TREDMPVCDVYVTLNCPHPETEHMINDETLKFKRGATVYNTARGLCDRDVAVALLES 300
 DB 241 TREDMPVCDVYVTLNCPHPETEHMINDETLKFKRGATVYNTARGLCDRDVAVALLES 300
 QY 301 GRLAGYAGDVWFPPQAPKDPHMTMPYNGMTPHISGTTTLTAQARYAAGTREILCEFEGR 360
 DB 301 GRLAGYAGDVWFPPQAPKDPHMTMPYNGMTPHISGTTTLTAQARYAAGTREILCEFEGR 360
 QY 361 PIRDEYLIYOGGALAGTAGHASYSGNATGSGEAAKFKKAV 401
 DB 361 PIRDEYLIYOGGALAGTAGHASYSGNATGSGEAAKFKKAV 401

RESULT 6
 AAU99110 ID AAU99110 standard; Protein; 401 AA.
 AC AAU99110;
 XX
 DT 24-SEP-2002 (first entry)
 DE Mycobacterium vaccae formate dehydrogenase C6S/C256a mutant.
 KM Formate dehydrogenase; alcohol production; ketone;
 XX mutant; mutein.
 OS Mycobacterium vaccae.
 OS Synthetic.
 XX

Key Location/Qualifiers
 FT Misc-difference 6 /note= "Wild type Cys substituted by Ser"
 FT Misc-difference 256 /note= "Wild type Cys substituted by Ala"
 FT
 XX
 PN EP1211316-AL.
 PD 05-JUN-2002.
 XX
 PE 27-NOV-2001; 2001EP-0128170.
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAI) DAICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from Mycobacterium
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX


```

KW mutant; mutein.
XX
OS Mycobacterium vaccae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 6
FT /note- "Wild type Cys substituted by Val"
FT 256
FT Misc-difference 256
FT /note- "Wild type Cys substituted by Ser"
XX
XX EP1211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
XX 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL ) DAICEL CHEM IND LTD.
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
XX
XX WPI; 2002-464925/50.
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
XX PT vaccae have enhanced activity in the presence of organic solvents and
XX PT are useful for producing alcohols from ketone raw materials
XX
XX Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
XX CC formate dehydrogenase protein which has strong formate dehydrogenase
XX CC activity in the presence of an organic solvent. The mutant formate
XX CC dehydrogenase protein of the invention may be used to produce alcohols
XX CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX CC proteins, the mutant polypeptides of the invention retain high enzymatic
XX CC activity in the presence of the alcohol product and so provides a higher
XX CC product yield. The present sequence represents a Mycobacterium vaccae
XX CC formate dehydrogenase protein mutant of the invention.
XX
XX Sequence 401 AA:
XX
Query Match 99.1%; Score 2115; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 1.7e-210;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

DB 361 PIRDEYLIYOGGALAGTAHYSKGNATGSEBAKFKRAV 401
|||||
RESULT 9
AAU09111
ID AAU09111 standard; Protein; 401 AA.
AC AAU09111;
XX
XX 24-SEP-2002 (first entry)
XX
XX Mycobacterium vaccae formate dehydrogenase C6S/C256V mutant.
XX DE
XX Formate dehydrogenase; alcohol production; ketone;
XX KM mutant; mutein.
XX
XX Mycobacterium vaccae.
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 6
XX FT /note- "Wild type Cys substituted by Ser"
XX FT 256
XX FT Misc-difference 256
XX /note- "Wild type Cys substituted by Val"
XX
XX EP1211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
XX 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL ) DAICEL CHEM IND LTD.
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
XX
XX WPI; 2002-464925/50.
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
XX PT vaccae have enhanced activity in the presence of organic solvents and
XX PT are useful for producing alcohols from ketone raw materials
XX
XX Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
XX CC formate dehydrogenase protein which has strong formate dehydrogenase
XX CC activity in the presence of an organic solvent. The mutant formate
XX CC dehydrogenase protein of the invention may be used to produce alcohols
XX CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX CC proteins, the mutant polypeptides of the invention retain high enzymatic
XX CC activity in the presence of the alcohol product and so provides a higher
XX CC product yield. The present sequence represents a Mycobacterium vaccae
XX CC formate dehydrogenase protein mutant of the invention.
XX
XX Sequence 401 AA:
XX
Query Match 99.1%; Score 2115; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 1.7e-210;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```


Db 121 AGIGSDHVDLSAIDRNVTAEVYTCNSISVAEHVMMILSVRNYLPSHEMARKGWN 180
Qy 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTWHA 240
Db 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTWHA 240
Qy 241 TREDMYPVCDDVYTLNCPILHPETEHMINDETCLKFRGAYIVNTARGKLCDDAVARALES 300
Db 241 TREDMYPVCDDVYTLNCPILHPETEHMINDETCLKFRGAYIVNTARGKLCDDAVARALES 300
Qy 301 GRLAGYAGDVWFPQAPKDHPRTPYNGMTPHISGTTLTQAQARYAACTREILCEFFEGR 360
Db 301 GRLAGYAGDVWFPQAPKDHPRTPYNGMTPHISGTTLTQAQARYAACTREILCEFFEGR 360
Qy 361 PIRDEYLLVGGALAGTGAHYSKGNATGSGSEAAKFKKAV 401
Db 361 PIRDEYLLVGGALAGTGAHYSKGNATGSGSEAAKFKKAV 401

RESULT 10
AAU99114

ID AAU99114 standard; Protein: 401 AA.

AC AAU99114;

DT 24-SEP-2002 (first entry)

DE Mycobacterium vaccae formate dehydrogenase C146S/C256S mutant.

KW Formate dehydrogenase; alcohol production; ketone;
mutant; mutein.OS Mycobacterium vaccae.
Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 146 /note= "Wild type Cys substituted by Ser"

FT Misc-difference 256 /note= "Wild type Cys substituted by Ser"

PN EPI211316-A1.

PD 05-JUN-2002.

PF 27-NOV-2001; 2001EP-0128170.

PR 29-NOV-2000; 2000JP-0363894.

PR 24-AUG-2001; 2001JP-0254631.

PA (DAIL) DAICEL CHEM IND LTD.

PI Mitsuhashi K, Yamamoto H, Kimoto N;

DR WPI; 2002-464925/50.

PT New mutant forms of formate dehydrogenase derived from Mycobacterium
vaccae have enhanced activity in the presence of organic solvents and
are useful for producing alcohols from ketone raw materials

PS Claim 8; Page -: 42pp; English.

XX This invention relates to a mutant form of the Mycobacterium vaccae
formate dehydrogenase protein which has strong formate dehydrogenase
activity in the presence of an organic solvent. The mutant formate
dehydrogenase protein of the invention may be used to produce alcohols
using ketones as raw materials. Unlike prior formatted dehydrogenase,
proteins, the mutant polypeptides of the invention retain high enzymatic
activity in the presence of the alcohol product and so provides a higher
product yield. The present sequence represents a Mycobacterium vaccae
formate dehydrogenase protein mutant of the invention.XX Sequence 401 AA;
SQ

Query Match 99.1%; Score 2115; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 1.7e-210;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAKVLCVYDDPVGVSFKTVARDDLPRKIDHYPGGILPTPAIDFTPGQLGSVSGELG 60
Db 1 MAKVLCVYDDPVGVSFKTVARDDLPRKIDHYPGGILPTPAIDFTPGQLGSVSGELG 60
Qy 61 RPYLESNGHTLVVTSDDKGPSPVFERELVDADVYISQFPWPAIYLPERIAKAKNLKALT 120
Db 61 RPYLESNGHTLVVTSDDKGPSPVFERELVDADVYISQFPWPAIYLPERIAKAKNLKALT 120
Qy 121 AGIGSDHVDLSAIDRNVTAEVYTCNSISVAEHVMMILSVRNYLPSHEMARKGWN 180
Db 121 AGIGSDHVDLSAIDRNVTAEVYTCNSISVAEHVMMILSVRNYLPSHEMARKGWN 180
Qy 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTWHA 240
Db 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTWHA 240
Qy 241 TREDMYPVCDDVYTLNCPILHPETEHMINDETCLKFRGAYIVNTARGKLCDDAVARALES 300
Db 241 TREDMYPVCDDVYTLNCPILHPETEHMINDETCLKFRGAYIVNTARGKLCDDAVARALES 300
Qy 301 GRLAGYAGDVWFPQAPKDHPRTPYNGMTPHISGTTLTQAQARYAACTREILCEFFEGR 360
Db 301 GRLAGYAGDVWFPQAPKDHPRTPYNGMTPHISGTTLTQAQARYAACTREILCEFFEGR 360
Qy 361 PIRDEYLLVGGALAGTGAHYSKGNATGSGSEAAKFKKAV 401
Db 361 PIRDEYLLVGGALAGTGAHYSKGNATGSGSEAAKFKKAV 401

RESULT 11
AAU99116

ID AAU99116 standard; Protein: 401 AA.

AC AAU99116;

DT 24-SEP-2002 (first entry)

DE Mycobacterium vaccae formate dehydrogenase C146S/C256V mutant.

KW Formate dehydrogenase; alcohol production; ketone;
mutant; mutein.OS Mycobacterium vaccae.
Synthetic.

FH Key Location/Qualifiers

FT Misc-difference 146 /note= "Wild type Cys substituted by Ser"

FT Misc-difference 256 /note= "Wild type Cys substituted by Val"

PN EPI211316-A1.

PD 05-JUN-2002.

PF 27-NOV-2001; 2001EP-0128170.

PR 29-NOV-2000; 2000JP-0363894.

PR 24-AUG-2001; 2001JP-0254631.

PA (DAIL) DAICEL CHEM IND LTD.

PI Mitsuhashi K, Yamamoto H, Kimoto N;

DR WPI; 2002-464925/50.

PT New mutant forms of formate dehydrogenase derived from Mycobacterium
vaccae have enhanced activity in the presence of organic solvents and

PT are useful for producing alcohols from ketone raw materials -
 XX Claim 8: Page -: 42pp: English.
 XX
 CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.
 CC
 XX
 SQ Sequence 401 AA:
 Query Match 99.1%; Score 2115; DB 23; Length 401;
 Best Local Similarity 99.5%; Pred. No. 1.7e-210;
 Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MAKVLCVLYDDPVYDGYPKTYARDDLPRKIDHYPGQILPFPKAIIDFTPGQLGVSSELGL 60
 DB 1 MAKVLCVLYDDPVYDGYPKTYARDDLPRKIDHYPGQILPFPKAIIDFTPGQLGVSSELGL 60
 QY 61 RPLIESNGHTLVYTSDDKGDPSVFERELVDADYVISOPEMPAYLTPERIKAKNKLALT 120
 DB 61 RPLIESNGHTLVYTSDDKGDPSVFERELVDADYVISOPEMPAYLTPERIKAKNKLALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWN 180
 DB 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWN 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTYDRHRLPESVEKELNTLWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTYDRHRLPESVEKELNTLWHA 240
 QY 241 TREDMPVCDVYVTLNCPLEHETHEMINDETLKLFRKRAYIVNTRARGLCRDVAARLES 300
 DB 241 TREDMPVCDVYVTLNCPLEHETHEMINDETLKLFRKRAYIVNTRARGLCRDVAARLES 300
 QY 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGALAGTAGHASYSGKNGATGSEEAKEKRAY 401
 DB 361 PIRDEYLIYOGALAGTAGHASYSGKNGATGSEEAKEKRAY 401
 RESULT 12
 AAU09119 standard; Protein; 401 AA.
 XX
 AC AAU09119;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DE *Mycobacterium vaccae* formate dehydrogenase C6a/C146a/C256v mutant.
 XX
 KM Formate dehydrogenase; alcohol production; ketone;
 KW mutant; muteln.
 XX
 OS *Mycobacterium vaccae*.
 OS
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 6 /note- "wild type Cys substituted by Ala"
 FT MISC-difference 146 /note- "wild type Cys substituted by Ala"
 FT MISC-difference 256 /note- "wild type Cys substituted by Val"
 XX

PN EP1211316-A1.
 XX
 PD 05-JUN-2002.
 XX
 PF 27-NOV-2001; 2001EP-0128170.
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAIIC) DAICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from *Mycobacterium*
 PT *vaccae* have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials -
 PS Claim 8; Page -: 42pp: English.
 XX
 CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.
 CC
 XX
 SQ Sequence 401 AA:
 Query Match 98.7%; Score 2107; DB 23; Length 401;
 Best Local Similarity 99.3%; Pred. No. 1.2e-209;
 Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MAKVLCVLYDDPVYDGYPKTYARDDLPRKIDHYPGQILPFPKAIIDFTPGQLGVSSELGL 60
 DB 1 MAKVLCVLYDDPVYDGYPKTYARDDLPRKIDHYPGQILPFPKAIIDFTPGQLGVSSELGL 60
 QY 61 RPLIESNGHTLVYTSDDKGDPSVFERELVDADYVISOPEMPAYLTPERIKAKNKLALT 120
 DB 61 RPLIESNGHTLVYTSDDKGDPSVFERELVDADYVISOPEMPAYLTPERIKAKNKLALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWN 180
 DB 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWN 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTYDRHRLPESVEKELNTLWHA 240
 DB 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTYDRHRLPESVEKELNTLWHA 240
 QY 241 TREDMPVCDVYVTLNCPLEHETHEMINDETLKLFRKRAYIVNTRARGLCRDVAARLES 300
 DB 241 TREDMPVCDVYVTLNCPLEHETHEMINDETLKLFRKRAYIVNTRARGLCRDVAARLES 300
 QY 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDPWRTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGR 360
 QY 361 PIRDEYLIYOGALAGTAGHASYSGKNGATGSEEAKEKRAY 401
 DB 361 PIRDEYLIYOGALAGTAGHASYSGKNGATGSEEAKEKRAY 401
 RESULT 13
 AAU09118 standard; Protein; 401 AA.
 XX
 AC AAU09118;
 XX
 DT 24-SEP-2002 (first entry)

XX Mycobacterium vaccae formate dehydrogenase C6A/C146S/C256V mutant.
DE Formate dehydrogenase; alcohol production; ketone;
XX mutant; mutEln.
KW
XX
OS Mycobacterium vaccae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild type Cys substituted by Ala"
FT Misc-difference 146 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
FT
XX
XX EPI211316-A1.
PN
XX
PD 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
PF
XX 29-NOV-2000; 2000JP-0363894.
PR 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL) DAICEL CHEM IND LTD.
PA
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
PI
DR WPI; 2002-464925/50.
XX
XX
PT New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
XX
PS Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
XX formate dehydrogenase protein which has strong formate dehydrogenase
XX activity in the presence of an organic solvent. The mutant formate
XX dehydrogenase protein of the invention may be used to produce alcohols
XX using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX proteins, the mutant polypeptides of the invention retain high enzymatic
XX activity in the presence of the alcohol product and so provides a higher
XX product yield. The present sequence represents a Mycobacterium vaccae
XX formate dehydrogenase protein mutant of the invention.
XX
SQ Sequence 401 AA;
Query Match 98.6%; Score 2106; DB 23; Length 401;
Best Local Similarity 99.3%; Pred. No. 1.5e-209;
Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVGVGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGYSGELG 60
DB 1 MAKVLCVLYDDPVGVGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGYSGELG 60
QY 61 RPYIESNGHLLVYTSDDKDPDSYFERELVDADVVISOPFMPAYLTPERIAKAKMLKIALT 120
DB 61 RPYIESNGHLLVYTSDDKDPDSYFERELVDADVVISOPFMPAYLTPERIAKAKMLKIALT 120
QY 121 AGTGSDDHVDLSAIDRNVTVAEVYYSNSISVAEHVMMILSLVNNYLPSEHMAKGGWNT 180
DB 121 AGTGSDDHVDLSAIDRNVTVAEVYYSNSISVAEHVMMILSLVNNYLPSEHMAKGGWNT 180
QY 181 ADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHTDRRLPESYEKELNTWHA 240
DB 181 ADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHTDRRLPESYEKELNTWHA 240
QY 241 TREDMYVCDVYVTLNCLPHPETEMINDETLKLFKRAYIVNTARGKLCORDAVARALE 300
DB 241 TREDMYVCDVYVTLNCLPHPETEMINDETLKLFKRAYIVNTARGKLCORDAVARALE 300

QY 301 GRLAGYAGDWFPQAPKDPHMPNMGMTPHISGTTLTNOARYAAGREILCFEFG 360
DB 301 GRLAGYAGDWFPQAPKDPHMPNMGMTPHISGTTLTNOARYAAGREILCFEFG 360
QY 361 PIRDEYLLVGGALAGTAGHSYSGKNATGSGSEAAKFKKAV 401
DB 361 PIRDEYLLVGGALAGTAGHSYSGKNATGSGSEAAKFKKAV 401
RESULT 14
AAU99107
ID AAU99107 standard; Protein: 401 AA.
XX
XX AAU99107;
XX
XX 24-SEP-2002 (first entry)
XX
XX
XX Mycobacterium vaccae formate dehydrogenase C6S/C146S/C256S mutant.
DE Formate dehydrogenase; alcohol production; ketone;
XX mutant; mutEln.
KW
XX
XX
OS Mycobacterium vaccae.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 146 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 256 /note= "Wild type Cys substituted by Ser"
FT
XX
XX EPI211316-A1.
PN
XX
XX
PD 05-JUN-2002.
XX
XX
PF 27-NOV-2001; 2001EP-0128170.
XX
XX
PR 29-NOV-2000; 2000JP-0363894.
PR 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL) DAICEL CHEM IND LTD.
PA
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
PI
DR WPI; 2002-464925/50.
XX
XX
PT New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
XX
PS Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
XX formate dehydrogenase protein which has strong formate dehydrogenase
XX activity in the presence of an organic solvent. The mutant formate
XX dehydrogenase protein of the invention may be used to produce alcohols
XX using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX proteins, the mutant polypeptides of the invention retain high enzymatic
XX activity in the presence of the alcohol product and so provides a higher
XX product yield. The present sequence represents a Mycobacterium vaccae
XX formate dehydrogenase protein mutant of the invention.
XX
SQ Sequence 401 AA;
Query Match 98.6%; Score 2105; DB 23; Length 401;
Best Local Similarity 99.3%; Pred. No. 1.9e-209;
Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVGVGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGYSGELG 60
DB 1 MAKVLCVLYDDPVGVGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGYSGELG 60

Db 1 MAKVLSVLYDDPVDGPKTYARDLPRKIDHPGQILPTPKAIDFTPGQLGSVSGELGL 60
 QY RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAYLTPERIARKAKNLKALT 120
 Db 61 RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAYLTPERIARKAKNLKALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWN 180
 Db 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWN 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPPDVLTHTDRRLPESYEKELNTWHA 240
 Db 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPPDVLTHTDRRLPESYEKELNTWHA 240
 QY 241 TREDMPVCDVVTLLNCPLEBETEMINDETLKLFKRCAYIVNTARGKLCRDVAVARALES 300
 Db 241 TREDMPVCDVVTLLNCPLEBETEMINDETLKLFKRCAYIVNTARGKLCRDVAVARALES 300
 QY 301 GRLAGYAGDVWFPQAPAKDHPWRTMPYNGMTPHISGTTTLTAQARYAAGTREILCEPFEGR 360
 Db 301 GRLAGYAGDVWFPQAPAKDHPWRTMPYNGMTPHISGTTTLTAQARYAAGTREILCEPFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401
 Db 361 PIRDEYLIYOGGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401

RESULT 15

AAU99120 standard; Protein; 401 AA.

AAU99120;

24-SEP-2002 (first entry)

Mycobacterium vaccae formate dehydrogenase C6S/C249S/C256V mutant.

Formate dehydrogenase; alcohol production; ketone;

mutant; muteln.

Mycobacterium vaccae.

Synthetic.

Key Location/Qualifiers

FT Misc-difference 6 /note= "wild type Cys substituted by Ser"

FT Misc-difference 249 /note= "wild type Cys substituted by Ser"

FT Misc-difference 256 /note= "wild type Cys substituted by Val"

FT Misc-difference 256 /note= "wild type Cys substituted by Val"

EP1211316-A1.

05-JUN-2002.

27-NOV-2001; 2001EP-0128170.

29-NOV-2000; 2000JP-0363894.

24-AUG-2001; 2001JP-0254631.

(DAIL) DAICEL CHEM IND LTD.

Mitsubishi K, Yamamoto H, Kimoto N;

WPI; 2002-464925/50.

New mutant forms of formate dehydrogenase derived from Mycobacterium
 vaccae have enhanced activity in the presence of organic solvents and
 are useful for producing alcohols from ketone raw materials

Example 5; Page -; 42pp; English.

This invention relates to a mutant form of the Mycobacterium vaccae

CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formated dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.

SO sequence 401 AA;

Query Match 98.6%; Score 2105; DB 23; Length 401;

Best Local Similarity 99.3%; Pred. No. 1.9e-209;

Matches 398; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAKVLSVLYDDPVDGPKTYARDLPRKIDHPGQILPTPKAIDFTPGQLGSVSGELGL 60
 Db 1 MAKVLSVLYDDPVDGPKTYARDLPRKIDHPGQILPTPKAIDFTPGQLGSVSGELGL 60
 QY 61 RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAYLTPERIARKAKNLKALT 120
 Db 61 RPYLESNGHTLVVTSKDGPDVSFERELVDADVVISQFPWPAYLTPERIARKAKNLKALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWN 180
 Db 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMIISLVNRYLPSEHMARKGWN 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPPDVLTHTDRRLPESYEKELNTWHA 240
 Db 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPPDVLTHTDRRLPESYEKELNTWHA 240
 QY 241 TREDMPVCDVVTLLNCPLEBETEMINDETLKLFKRCAYIVNTARGKLCRDVAVARALES 300
 Db 241 TREDMPVCDVVTLLNCPLEBETEMINDETLKLFKRCAYIVNTARGKLCRDVAVARALES 300
 QY 301 GRLAGYAGDVWFPQAPAKDHPWRTMPYNGMTPHISGTTTLTAQARYAAGTREILCEPFEGR 360
 Db 301 GRLAGYAGDVWFPQAPAKDHPWRTMPYNGMTPHISGTTTLTAQARYAAGTREILCEPFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401
 Db 361 PIRDEYLIYOGGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401

Search completed: July 24, 2003, 20:46:01

Job time : 93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:44:08 ; Search time 42 Seconds

(without alignments)
918.182 Million cell updates/sec

Title: US-09-996-008b-2

Perfect score: 2135

Sequence: 1 MAKVLCVLYDDPDVGYPKTY.....YSKGNATGSEPAKKFKAAY 401

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2036	95.4	393	2	JU0334
2	1806	84.6	400	2	UC7815
3	1704	79.8	401	2	C95293
4	910.5	42.6	375	2	A47117
5	883.5	41.4	377	2	S30088
6	872	40.8	384	2	T51423
7	868	40.7	379	2	J02272
8	812	38.0	364	2	UC4252
9	803	37.6	376	2	S67300
10	777	36.4	374	2	E89779
11	436	20.4	236	2	S65308
12	373	17.5	326	2	C72483
13	352	16.5	333	2	D75067
14	348	16.3	525	2	H69229
15	347	16.3	145	2	S65309
16	346.5	16.2	333	2	C71165
17	340	15.9	354	2	AB3200
18	340	15.8	525	1	C69705
19	338	15.8	335	2	B75057
20	337.5	15.8	307	2	E71011
21	336.5	15.6	323	2	B69472
22	332	15.6	328	2	H95353
23	332	15.3	376	2	A71175
24	326	15.2	524	2	A64427
25	323.5	14.9	322	2	S48189
26	319	14.6	540	2	B83850
27	312.5	14.6	534	2	H89956
28	311.5	14.4	360	2	AB3496
29	308				

30	305.5	14.3	334	2	AH2585	hypothetical prote
31	305.5	14.3	334	2	G97367	probable dehydroge
32	299	14.0	327	2	A72390	hypothetical prote
33	298.5	14.0	529	2	T35831	probable D-3-phosp
34	297	13.9	410	1	C64070	D-3-phosphoglycerate d
35	296.5	13.9	531	2	A98272	D-3-phosphoglycerate d
36	296.5	13.9	531	2	AF3012	D-3-phosphoglycerate d
37	296	13.9	318	2	G96910	D-3-phosphoglycerate d
38	293.5	13.7	526	2	E87647	D-3-phosphoglycerate d
39	292.5	13.7	533	2	A70464	D-3-phosphoglycerate d
40	290	13.6	527	2	E69351	D-3-phosphoglycerate d
41	289	13.5	334	2	G70363	D-lactate dehydrog
42	288	13.5	318	2	AG1084	phosphoglycerate d
43	287	13.4	442	2	H87710	hypothetical prote
44	287	13.4	527	2	C84393	phosphoglycerate d
45	285.5	13.4	412	2	AD3611	phosphoglycerate d

ALIGNMENTS

RESULT 1

JU0334

formate dehydrogenase (EC 1.2.1.2) - Pseudomonas sp.

N:Alternate names: NAD-dependent formate dehydrogenase

C:Species: Pseudomonas sp.

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 23-Sep-2002

C:Accession: JU0334

R:Popov, V.O.; Shumilina, I.A.; Ustinikova, T.B.; Lamzin, V.S.; Egorov, T.A.

Bioorg. Khim. 16, 324-335, 1990

A:Title: NAD-dependent formate dehydrogenase from methylotrophic bacterium Pseudomona

A:Reference number: JU0334; MUID:90290536; PMID:2357236

A:Accession: JU0334

A:Molecule type: protein

A:Residues: 1-393 <POP>

A>Note: article in Russian with English abstract

C:Superfamily: Neurospora formate dehydrogenase

C:Keywords: homodimer; NAD; oxidoreductase

Query Match 95.4%; Score 2036; DB 2; Length 393;

Best local similarity 97.7%; Pred. No. 3e-161;

Matches 384; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY	2	AKVLCVLYDDPDVGYPKTYARDLPRKIDHYGGQTLPRKAIIDTFPGQLLSVSGELGLR	61
DB	1	AKVLCVLYDDPDVGYPKTYARDLPRKIDHYGGQTLPRKAIIDTFPGQLLSVSGELGLR	60
QY	62	PYLESNGHTLVVTSKDGSDSVFERELVDADVISQPPMPAYLTPERTAKAKNLKALTA	121
DB	61	KYLESNGHTLVVTSKDGSDSVFERELVDADVISQPPMPAYLTPERTAKAKNLKALTA	120
QY	122	GIGSDHVDLQSAIDRNVTVAETVYCNISVAEHVVMILSLVRNLYPSHEMARKGWMA	181
DB	121	GIGSDHVDLQSAIDRNVTVAETVYCNISVAEHVVMILSLVRNLYPSHEMARKGWMA	180
QY	182	DCVSHAVDLEAMHGTVAAGRIGLAVALRRLRPFQVHLHYTRNHRPESVEKELNTWTAT	241
DB	181	DCVSHAVDLEAMHGTVAAGRIGLAVALRRLRPFQVHLHYTRNHRPESVEKELNTWTAT	240
QY	242	REDMPVCDVVTTLNCPLEPTEHMIINDETLFLFKGATIVNTARGKLCDDRAVARALESG	301
DB	241	REDMPVCDVVTTLNCPLEPTEHMIINDETLFLFKGATIVNTARGKLCDDRAVARALESG	300
QY	302	RLAGYAGDVPFPPAPKDPHPRMPYNGMTPHISGTTVLTQAARYAAGTRELIECFEGRP	361
DB	301	RLAGYAGDVPFPPAPKDPHPRMPYNGMTPHISGTTVLTQAARYAAGTRELIECFEGRP	360
QY	362	IRDEYLVYOGGALAGTGAHSTSGKNATGSEEA	394
DB	361	IRDEYLVYOGGALAGTGAHSTSGKNATGSEEA	393

RESULT 2

JC7815
 formate dehydrogenase (EC 1.2.1.2) - *Paracoccus* sp. (Strain 12-A)
 C:Species: *Paracoccus* sp.
 C>Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
 C:Accession: JC7815
 R:Shinoda, T.; Satoh, T.; Mineki, S.; Iida, M.; Taguchi, H.
 Biosci. Biotechnol. Biochem. 66, 271-276, 2002
 A:Title: Cloning, nucleotide sequencing, and expression in *Escherichia coli* of the gene
 A:Reference number: JC7815; PMID:1199398; MUID:21994041
 A:Accession: JC7815
 A:Molecule type: DNA
 A:Residues: 1400 <SHD>
 A:Cross-references: DDBJ:AB071373
 C:Comment: This enzyme, which catalyzes the conversion of formate to carbon dioxide with
 Li-enzyme systems, such as bioreactors involving NADH as a coenzyme, and plays a key role
 C:Genetics:
 A:Gene: fdh
 C:Keywords: oxidoreductase

Query Match 84.6%; Score 1806; DB 2; Length 400;
 Best Local Similarity 83.2%; Pred. No. 3.9e-142;
 Matches 333; Conservative 28; Mismatches 39; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTFGQLGSGVSGELGL 60
 D 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTFGQLGSGVSGELGL 60
 QY 61 RPYLESNGHTLVVTSDDKGPDSVFERELVDADVVISOPFMPAYLTPERIAKAKNLKALT 120
 D 61 RNYLEAGHGLVYVTSDDKGPDSVFERELVDADVVISOPFMPAYLTPERIAKAKNLKALT 120
 QY 121 AGISDHDVDDQSAIDRNVTYAEVYTCNSISVAEHVMMILSLVKNYLPSEHARKGWN 180
 D 121 AGISDHDVDDQSAIDRNVTYAEVYTCNSISVAEHVMMILSLVKNYLPSEHARKGWN 180
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVLTHTDRRLPESVEKELNTWHA 240
 D 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVLTHTDRRLPESVEKELNTWHA 240
 QY 241 TREDMYPVCVVTINCLHPETEHEMINDETLKLFRKAGYIVNTARGKLCRDVAARALE 300
 D 241 TREDMYPVCVVTINCLHPETEHEMINDETLKLFRKAGYIVNTARGKLCRDVAARALE 300
 QY 301 GRLAGYAGDWFPPAPAKDHPWRTMPYNGMTPHISGTTLTAAQRYAAGTRELIECFEGR 360
 D 301 GRLAGYAGDWFPPAPAKDHPWRTMPYNGMTPHISGTTLTAAQRYAAGTRELIECFEGR 360
 QY 361 PIRDEVLIYOGGALAGAGAHSGKNGATGSGSEAKRKA 400
 D 361 PIRDEVLIYOGGALAGAGAHSGKNGATGSGSEAKRKA 400

RESULT 3
 C95293
 Probable NAD-dependent formate dehydrogenase [Imported] - *Sinorhizobium meliloti* (strain
 C:Species: *Sinorhizobium meliloti*
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 23-Sep-2002
 C:Accession: C95293
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
 ; Kalman, S.; Keating, D.H.; Palm, C.; Beck, M.C.; Surzycki, R.; Wells, D.H.; Yen, K.C.;
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti*
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: C95293
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1401 <KUR>
 A:Cross-references: GB:AE006469; PIDN:AAK64909.1; PID:914523329; GSPDB:GN00165
 R:Experimental sources: strain 1021, megaplasmid pSymA
 R:Galbert, F.; Flinn, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
 heblut, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen,
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Sma0478
 A:Genome: Plasmid
 C:Superfamily: Neurospora formate dehydrogenase

Query Match 79.8%; Score 1704; DB 2; Length 401;
 Best Local Similarity 78.4%; Pred. No. 1.2e-133;
 Matches 312; Conservative 39; Mismatches 47; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTFGQLGSGVSGELGL 60
 D 3 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTFGQLGSGVSGELGL 62
 QY 61 RPYLESNGHTLVVTSDDKGPDSVFERELVDADVVISOPFMPAYLTPERIAKAKNLKALT 120
 D 63 RKFLEGGHTLVVTSDDKGPDSVFERELVDADVVISOPFMPAYLTPERIAKAKNLKALT 122
 QY 121 AGISDHDVDDQSAIDRNVTYAEVYTCNSISVAEHVMMILSLVKNYLPSEHARKGWN 180
 D 123 AGISDHDVDDQSAIDRNVTYAEVYTCNSISVAEHVMMILSLVKNYLPSEHARKGWN 182
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVLTHTDRRLPESVEKELNTWHA 240
 D 183 ADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVLTHTDRRLPESVEKELNTWHA 242
 QY 241 TREDMYPVCVVTINCLHPETEHEMINDETLKLFRKAGYIVNTARGKLCRDVAARALE 300
 D 243 TAEAMVPCVVTINCLHPETEHEMINDETLKLFRKAGYIVNTARGKLCRDVAARALE 302
 QY 301 GRLAGYAGDWFPPAPAKDHPWRTMPYNGMTPHISGTTLTAAQRYAAGTRELIECFEGR 360
 D 303 GRLAGYAGDWFPPAPAKDHPWRTMPYNGMTPHISGTTLTAAQRYAAGTRELIECFEGR 362
 QY 361 PIRDEVLIYOGGALAGAGAHSGKNGATGSGSEAKRKA 398
 D 363 PIRDEVLIYOGGALAGAGAHSGKNGATGSGSEAKRKA 400

RESULT 4
 A47117
 formate dehydrogenase (EC 1.2.1.2) - *Neurospora crassa*
 C:Species: *Neurospora crassa*
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 23-Sep-2002
 C:Accession: A47117
 R:Chow, C.M.; RajBhandary, U.L.
 J. Bacteriol. 175, 3703-3709, 1993
 A:Title: Developmental regulation of the gene for formate dehydrogenase in *Neurospora*
 A:Reference number: A47117; MUID:93265982; PMID:8509325
 A:Accession: A47117
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-375 <CHO>
 A:Cross-references: GB:U13964; NID:g293949; PIDN:AAA99900.1; PID:g1321604
 C:Genetics:
 A:Insertions: 1/3; 17/3
 C:Superfamily: Neurospora formate dehydrogenase
 C:Keywords: oxidoreductase

Query Match 42.6%; Score 910.5; DB 2; Length 375;
 Best Local Similarity 48.4%; Pred. No. 8.8e-68;
 Matches 196; Conservative 56; Mismatches 112; Indels 41; Gaps 8;

QY 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTFGQLGSGVSGELGL 60
 D 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGQILPTPKAIDFTFGQLGSGVSGELGL 31
 QY 61 RPYLESNGHTLVVTSDDKGPDSVFERELVDADVVISOPFMPAYLTPERIAKAKNLKALT 120
 D 61 RPYLESNGHTLVVTSDDKGPDSVFERELVDADVVISOPFMPAYLTPERIAKAKNLKALT 120

Db 32 RKLWEDGHTLVTTCKDGENSTFDEKLEDAEIIITTPFHPCGYLTAERLARAKKIKLAVT 91
 QY 121 AGISGDHVDLQSAIDRN--VTVAEVYTCNSISVAEHVVMILSLVRNLTSPSEHARKGCM 178
 Db 92 AGISGDHVDLNAKNTNGITVAEYTCNSISVAEHVVMILSLVRNLTSPSEHARKGCM 151
 QY 179 NIADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDV-HLHYDRHRLPSEVKEKELMT 237
 Db 152 DVAAKAKNEFDLEKGVGTGCVGRIGERVLRLKPFCKELLYDYQGLSAEKEKEICGR 211
 QY 238 WHATREDMYPCVDVYTLNCPRLPETHHINDETLLKFRGAYIVTAGKLCDRDAVARA 297
 Db 212 RVADLEMLACDQVYTLNCPRLPETHHINDETLLKFRGAYIVTAGKLCDRDAVARA 271
 QY 298 LESGRLAGYAGDVWFPAPKDPMP--TMPY---NGMTPHISGTLTAQARYAAGTREI 352
 Db 272 LKSGHLKRGYGDVWFPAPKDPMP--TMPY---NGMTPHISGTLTAQARYAAGTREI 331
 QY 353 LECFEFGR-PIRDEYLVVQGALAGTGAHSYSKGNATGKSEBAK 396
 Db .332 IESYLSGKHQYRPEDLYVGGDYA---TKSYGERERAKAAAAA 373

RESULT 5

S30088
 acia protein - Emeritella nidulans
 C:Species: Emeritella nidulans, Aspergillus nidulans
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Sep-2002
 C:Accession: S30088
 R:Saleeba, J.A.; Cobbett, C.S.; Hynes, M.J.
 Mol. Gen. Genet. 235, 349-358, 1992
 A:Title: Characterization of the anda-regulated acia gene of Aspergillus nidulans.
 A:Reference number: S30088; MUID:93101140; PMID:1465107
 A:Accession: S30088
 A:Molecule type: DNA
 A:Residues: 1-377 <SAL>
 A:Cross-references: EMBL:211612; NID:95544; PIDN:CAA7687.1; PID:95545
 C:Genetics:
 A:Gene: acia
 A:Map position: 1
 A:Introns: 5/3; 64/2
 C:Superfamily: Neurospora formate dehydrogenase

Query Match 41.4%; Score 883.5; DB 2; Length 377;
 Best Local Similarity 52.1%; Pred. No. 1.6e-65;
 Matches 185; Conservative 48; Mismatches 95; Indels 27; Gaps 6;
 QY 47 PGQLGSVSGELGRLPYLESNGHTLVTSKDGPDSEFERELVDADVVISQFMPAYLTP 106
 Db 14 PG-LGTETNELGIRKWTIEQGHVLTSSDKDGENSTFDEKLEVDVAVITTPFHGYLTA 72
 QY 107 ERIKAKKILKLTAGISGDHVDLQSAIDRN--VTVAEVYTCNSISVAEHVVMILSLVR 164
 Db 73 ERLKAKKILKLTAVAGISGDHVDLQSAIDRN--VTVAEVYTCNSISVAEHVVMILSLVR 132
 QY 165 NYLPSHEWARKGNINADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDV-HLHYDR 223
 Db 133 NFVFAHQDIRGDMNVAAVAKNEFDLENKVVGTGCVGRIGERVLRLKPFCKELLYDY 192
 QY 224 HRLPESVEKELNLTWHATREDMYPCVDVYTLNCPRLPETHHINDETLLKFRGAYIVTAG 272
 Db 193 QPLPEVEKEIGARVSLSEMVSSQDVYTLNCPRLPETHHINDETLLKFRGAYIVTAG 252
 QY 273 -----LEKRGAYIVTAGKLCDRDAVARALESGRLAGYAGDVWFPAPKDPMP 326
 Db 253 LIIPMLMTHKOSWLVNARGALVYKEDVAELKSGHLRGYGDVWFPAPKDPMP 312
 QY 327 Y-----NGMTPHISGTLTAQARYAAGTREI LECFEFGR-PIRDEYLVVQGALA 375
 Db 313 HPWGGKATVPHMSGTSLAQLIRYANGKAILDSYFSRPFVQPDLLYVHGDDYA 367

RESULT 6

TS1423
 formate dehydrogenase (FDH) - Arabidopsis thaliana
 N:Alternate names: protein T9L3 80
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 23-Sep-2002
 C:Accession: TS1423
 R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.;
 submitted to the Protein Sequence Database, August 2000
 A:Reference number: 225394
 A:Accession: TS1423
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-384 <SAT>
 A:Cross-references: EMBL:AL391149
 A:Experimental source: cultivar Columbia; BAC clone T9L3
 C:Genetics:
 A:Map position: 5
 A:Introns: 29/3; 88/2; 347/3
 A:Note: T9L3_80
 C:Superfamily: Neurospora formate dehydrogenase

Query Match 40.8%; Score 872; DB 2; Length 384;
 Best Local Similarity 52.9%; Pred. No. 1.4e-64;
 Matches 172; Conservative 44; Mismatches 109; Indels 0; Gaps 0;

QY 51 LGSVGEGLRPLRYLESNGHTLVTSKDGPDSEFERELVDADVVISQFMPAYLTP 110
 Db 56 LGSVGEGLRPLRYLESNGHTLVTSKDGPDSEFERELVDADVVISQFMPAYLTP 115
 QY 111 KAKNLKLTAGISGDHVDLQSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNLTSP 170
 Db 116 KAKNLKLTAGISGDHVDLQSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNLTSP 175
 QY 171 EMARKGNNIADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYDRHRLPESV 230
 Db 176 NOVYAGENNVAIAVRAVDLEKGTGTGAGAGIGLTLQRLKPFCCNLTLDRLQMAPEL 235
 QY 231 EKELNLTWHATREDMYPCVDVYTLNCPRLPETHHINDETLLKFRGAYIVTAGKLC 290
 Db 236 EKELNLTWHATREDMYPCVDVYTLNCPRLPETHHINDETLLKFRGAYIVTAGKLC 295
 QY 291 RDAVARALESGRLAGYAGDVWFPAPKDPMPRTMPYNGMTPHISGTLTAQARYAAGT 350
 Db 296 RDAVARALESGRLAGYAGDVWFPAPKDPMPRTMPYNGMTPHISGTLTAQARYAAGT 355
 QY 351 ELLECFFGRPIRDEYLVVQGALA 375
 Db 356 DMLERYFKGEDEPTEYNTYKDELA 380

RESULT 7

JQ2272
 formate dehydrogenase (EC 1.2.1.2) precursor, mitochondrial - potato
 C:Species: Solanum tuberosum (potato)
 C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 23-Sep-2002
 C:Accession: JQ2272; PC2189; S33125
 R:Des Francs-Small, C.C.; Ambard-Bretteville, F.; Small, I.D.; Remy, R.
 Plant Physiol. 102, 1171-1177, 1993
 A:Title: Identification of a major soluble protein in mitochondria from nonphotosynth
 A:Reference number: JQ2272; MUID:94105343; PMID:8278546
 A:Accession: JQ2272
 A:Molecule type: mRNA
 A:Residues: 1-379 <DESI>
 A:Cross-references: EMBL:Z21493; NID:q297797; PIDN:CAA79702.1; PID:q297798
 A:Accession: PC2189
 A:Molecule type: protein
 A:Residues: 24-50 <DESI>
 C:Superfamily: Neurospora formate dehydrogenase
 C:Keywords: mitochondrion; oxidoreductase
 F:1-23/Domain: transit peptide (mitochondrion) #status predicted <TM>
 F:24-379/Product: formate dehydrogenase #status predicted <Mat>
 F:199-222/Region: NAD binding #status predicted

QY 350 RE-ILECFEGR 361
 Db 306 IENIVRLKGGK 318

RESULT 13

Probable dehydrogenase PAB1008 - Pyrococcus abyssi (strain Orsay)
 C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: D75067
 R:anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
 A:Reference number: A75001
 A:Accession: D75067
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-333 <KAM>
 A:Cross-references: GB:AJ248287; GB:AL096836; NID:95458657; PIDN:CA85043.1; PID:9545894
 A:Experimental source: strain Orsay
 C:Genetics:
 A:Gene: PAB1008
 C:Superfamily: phosphoglycerate dehydrogenase

Query Match 16.5% Score 352; DB 2; Length 333;
 Best Local Similarity 32.3%; Pred. No. 1.8e-21;
 Matches 94; Conservative 60; Mismatches 105; Indels 32; Gaps 12;

QY 61 RPYLESNGHTLVTSKDGSDVPERLY-DADVVISQPFMPVATLPERIAKANLKLAL 119
 Db 14 KPLEELKKYDVL--KPYSEELKEIPEDDIIAPV--TRITDILERRARLVIS 69
 QY 120 TAGIGSDHVDLSAIDRNVTVAEYTCNSISVAEHVMMILSLVRYNLPSEHARKGWN 179
 Db 70 QDSAGYHVDVEATKRGIVTVKSGLSSEVAFALGLISLMRKHYADSFREGKE 129
 QY 180 IADCVSHAY-----DLEAMH--VGTVAAGRIGLAVLRRIAPDVHLHTDRHRESEV 230
 Db 130 -----SHTFWMRERKEVELYKGEVGIYGMALGKAIARLRKPGCEIYWSRRR-KEDI 183
 QY 231 EKEINLWMTNREMYVVCVVTNCPHPEHEMINDETLKLFRGATVNTARKLGD 290
 Db 184 EREVNNAK-YDLDDELLEVDIVILLPLTKETTHIINEEVKKLE-GKYLVINIGRALID 241
 QY 291 RDVVARALIESGRLAGYADVFPQAPAKDH-----PWRTMPYNGMPHISG 336
 Db 242 EKMLVKAIRKGLKGFATDVFEERP-KEHELRFRKRYV---LTTHYAG 287

RESULT 14

phosphoglycerate dehydrogenase - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 11-Jun-1999
 C:Accession: H69229
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qi, D.; Spadafora, R.; Viscardi, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivani, N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
 A:Reference number: A69000; M01D:96037514; PMID:9371463
 A:Accession: H69229
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-525 <MTH>
 A:Cross-references: GB:AE000870; GB:AE000666; NID:92622039; PIDN:AB85466.1; PID:9262206
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH970
 C:Superfamily: Bacillus phosphoglycerate dehydrogenase

Query Match 16.3% Score 348; DB 2; Length 525;

Best Local Similarity 31.7%; Pred. No. 7.6e-21;
 Matches 97; Conservative 54; Mismatches 131; Indels 24; Gaps 6;

QY 72 VVTSKDGSDPSVFERELVDADVVISQPFMPAYL-----TPERIAKAK 113
 Db 6 VLINDSINEKISLEEV-REVVTNTITPEELDAIKDDALVRSRTKTRVIEAAP 64
 QY 114 NKLALTAGIGSDHVDLSAIDRNVTVAEYTCNSISVAEHVMMILSLVRYNLPSEHWA 173
 Db 65 RLKTIARAGYGVNDVKAATDRGIMVINAPESSISVAEHSISLMALARKIADRSV 124
 QY 174 RKGWNTADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRIAPDVHLHTDRHRESEVEKE 233
 Db 125 KEKWEENRRMP--GIELNGKTLIGIGRIGSQVYVETKAFGMDIMVYDYSKEAAE-E 181
 QY 234 LNLWMTNREMYVVCVVTNCPHPEHEMINDETLKLFRGATVNTARKLGD 293
 Db 182 MGVY-VTDLETLRESQIVTTHVPLTPETRLISEDFKLMKDTAFVNCARGGIDEDA 240
 QY 294 VARALIESGRLAGYADVFPQAPAKDHWPRTMPYNGMPHISGTTTQAARYAGTREIL 353
 Db 241 LYRLKDGELAGALDVFEEEP-PEGSPLELENVVLTPLHGASTSAQDDAALIVANEI 299
 QY 354 ECFPEG 359
 Db 300 KTVFQG 305

RESULT 15

S65309
 Probable membrane protein YPL276w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein P0323
 C:Species: Saccharomyces cerevisiae
 C:Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 23-Sep-2002
 C:Accession: S65309
 R:Duisterhoef, A.; Floeth, M.; Filtz, M.; Hilbert, H.; Moestl, D.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S65292
 A:Accession: S65309
 A:Molecule type: DNA
 A:Residues: 1-145 <DUE>
 A:Cross-references: EMBL:Z73632; NID:91370566; PIDN:CAA98012.1; PID:91370567; MIPS:YP
 A:Experimental source: strain S288C (AB972)
 R:Delius, H.; Hebling, U.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64967
 A:Accession: S65330
 A:Molecule type: DNA
 A:Residues: 1-145
 A:Cross-references: EMBL:Z73632; NID:91370566; PIDN:CAA98012.1; PID:91370567; MIPS:YP
 A:Experimental source: strain S288C (AB972)
 C:Genetics:
 A:Cross-references: SGD:S0006197
 A:Map position: 16L
 A:Superfamily: Neurospora formate dehydrogenase
 C:Keywords: transmembrane protein
 F:121-137/Domain: transmembrane #status predicted <TM>

Query Match 16.3% Score 347; DB 2; Length 145;
 Best Local Similarity 44.7%; Pred. No. 1.4e-21;
 Matches 76; Conservative 27; Mismatches 37; Indels 30; Gaps 3;

QY 3 KVLVLALDDPDVDGPKTYARDLPKIDHYFGGQILPPKAIIDFTPGGLGSVSGELRLP 62
 Db 5 KVLVLALD-----GKHAHQE-----KLDCIENEGIRN 35
 QY 63 YLESNGHTLVTSKDG-SPDSVFERELVDADVVISQPFMPAYLTPERIAKANLKLALTA 121
 Db 36 FIEOGYELVTTIDKDEPTSTVDRELKDAEIVTTFPPAYISRNRIAPAPMLKLCVTA 95
 QY 122 GIGSDHVDLSAIDRNVTVAEYTCNSISVAEHVMMILSLVRYNLPSEH 171
 Db 96 GVGSDHVDLEAMERKITVETVGSNVVSAEHMATIIVLIRYNGHQ 145

Fri Jul 25 10:38:27 2003

us-09-996-008b-2_1.rpr

Page 7

Search completed: July 24, 2003, 20:49:08
Job time : 43 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:49:14 ; Search time 83 Seconds
(without alignments)

766.860 Million cell updates/sec

Title: US-09-996-008b-2

Perfect score: 401

Sequence: 1 MAKVLCVLDDPDVGYPKTY.....YSKGNATGSGSEAKFKKAV 401

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_19Jun03:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	401	100.0	401	23	AAU99122
2	300	74.8	401	23	AAU99112
3	300	74.8	401	23	AAU99113
4	300	74.8	401	23	AAU99115
5	294	73.3	401	23	AAU99108
6	294	73.3	401	23	AAU99109
7	294	73.3	401	23	AAU99110
8	294	73.3	401	23	AAU99111
9	294	73.3	401	23	AAU99117

10	249	62.1	401	23	AAU99121	Mycobacterium vacc	
11	242	60.3	401	23	AAU99120	Mycobacterium vacc	
12	223	55.6	401	19	AAH37857	Formate dehydrogen	
13	199	49.6	401	23	AAU99114	Mycobacterium vacc	
14	199	49.6	401	23	AAU99116	Mycobacterium vacc	
15	193	48.1	401	23	AAU99107	Mycobacterium vacc	
16	193	48.1	401	23	AAU99118	Mycobacterium vacc	
17	193	48.1	401	23	AAU99119	Mycobacterium vacc	
18	192	47.9	401	23	AAU99121	Methylophilic pse	
19	29	7.2	401	23	ABH77565	Ancyclobacter form	
20	20	5.0	399	21	AAU85650	Formate dehydrogen	
21	21	16	4.0	362	10	AAU94763	Formate dehydroge
22	16	4.0	364	18	AAW16414	Candida boidinii s	
23	16	4.0	365	22	AAU63573	Amino acid sequenc	
24	16	4.0	365	22	AAU63574	Synthetic amino ac	
25	11	2.7	376	22	AAU60245	Rice formate dehyd	
26	11	2.7	384	23	ABH83475	Hericidially activ	
27	10	2.5	10	21	AAU85652	Formate dehydrogen	
28	9	2.2	319	21	AAU64710	Arabidopsis thailia	
29	9	2.2	321	21	AAU64709	Arabidopsis thailia	
30	9	2.2	340	21	AAU64708	Arabidopsis thailia	
31	8	2.0	20	21	AAU85651	Formate dehydrogen	
32	8	2.0	190	22	ABH64123	Drosophila melanog	
33	8	2.0	253	20	AAU04477	Rhodococcus rhodoc	
34	8	2.0	335	22	AAU34334	Staphylococcus aur	
35	8	2.0	343	22	AAU37388	Staphylococcus aur	
36	8	2.0	343	22	AAU37557	Staphylococcus aur	
37	8	2.0	398	22	AAU81109	Mycobacterium tuben	
38	8	2.0	550	18	AAU31585	Streptococcus pneu	
39	8	2.0	550	20	AAU42391	A novel PROs amino	
40	8	2.0	612	23	ABP30503	Streptococcus poly	
41	8	2.0	617	22	AAU37635	Streptococcus pneu	
42	8	2.0	617	23	ABP26564	Streptococcus poly	
43	8	2.0	617	24	ABU00611	S. pneumoniae type	
44	8	2.0	618	20	AAU88486	Streptococcus pneu	
45	8	2.0	618	23	ABP26565	Streptococcus poly	

ALIGNMENTS

RESULT 1

AAU99122 standard; Protein: 401 AA.

AAU99122:

24-SEP-2002 (first entry)

Mycobacterium vaccae formate dehydrogenase protein.

Formate dehydrogenase; alcohol production; ketone; mutant.

Mycobacterium vaccae.

EP1211316-A1.

05-JUN-2002.

27-NOV-2001; 2001EP-0128170.

29-NOV-2000; 2000JP-0363894.

24-AUG-2001; 2001JP-0254631.

(DAI) DAICEL CHEM IND LTD.

Mitsunashi K, Yamamoto H, Kimoto N.

WPI: 2002-464925/50.

N-PSDB: ABK86863.

New mutant forms of formate dehydrogenase derived from Mycobacterium

PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials -
 XX
 PS Example 1; Page 23-24; 42pp; English.

CC This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents the Mycobacterium vaccae
 CC formate dehydrogenase protein of the invention. This sequence was used
 CC to generate the mutant proteins of the invention.

XX Sequence 401 AA:

Query Match 100.0%; Score 401; DB 23; Length 401;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPVDGPKTKYARDLPRKIDHYRGGQLTPPKAIDFTPGQLLSVSGELG 60
 DB 1 MAKVLCVLYDDPVDGPKTKYARDLPRKIDHYRGGQLTPPKAIDFTPGQLLSVSGELG 60
 QY 61 RPYLESNGHTLVVTSKDDPDSYFERELVDADVISQPPMPAYLTPPERIAKAKNKLALT 120
 DB 61 RPYLESNGHTLVVTSKDDPDSYFERELVDADVISQPPMPAYLTPPERIAKAKNKLALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEYTYCNSISVAEHVVMILSLVRNLPSEHARRKGMNI 180
 DB 121 AGIGSDHVDLQSAIDRNVTVAEYTYCNSISVAEHVVMILSLVRNLPSEHARRKGMNI 180
 QY 181 ADCVSHAUDLEAMHVTGAAGRIGLAVLRRLAFDVLHLYTDRHRLPESVEKELNLTWHA 240
 DB 181 ADCVSHAUDLEAMHVTGAAGRIGLAVLRRLAFDVLHLYTDRHRLPESVEKELNLTWHA 240
 QY 241 TREDMYPCVDVYTLNCLPRLPETEHMINDETLLKFKRGAYIVNTARGKLCDDRAVARALES 300
 DB 241 TREDMYPCVDVYTLNCLPRLPETEHMINDETLLKFKRGAYIVNTARGKLCDDRAVARALES 300
 QY 301 GRLAGYAGDVWFPQAPKDPWMTMPYNGMTPHISGTLTQAARYAAGTREILCEFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDPWMTMPYNGMTPHISGTLTQAARYAAGTREILCEFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSEBAKFKKAV 401
 DB 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSEBAKFKKAV 401

RESULT 2
 AAU99112
 ID AAU99112 standard; Protein; 401 AA.

XX AAU99112;

DT 24-SEP-2002 (first entry)

DE Mycobacterium vaccae formate dehydrogenase C146S mutant;

XX Formate dehydrogenase; alcohol production; ketone;

KW mutant; mutelin.

XX Mycobacterium vaccae.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 146 /note= "Wild type Cys substituted by Ser"

XX EP1211316-A1.

PD 05-JUN-2002.

XX 27-NOV-2001; 2001EP-0128170.

XX 29-NOV-2000; 2000JP-0363894.

PR 24-AUG-2001; 2001JP-0254631.

PA (DAI) DAICEL CHEM IND LTD.

PI Mitsuhashi K, Yamamoto H, Kimoto N;

DR WPI; 2002-464925/50.

PT New mutant forms of formate dehydrogenase derived from Mycobacterium

PT vaccae have enhanced activity in the presence of organic solvents and

PT are useful for producing alcohols from ketone raw materials -

XX Claim 8; Page -: 42pp; English.

CC This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.

XX Sequence 401 AA:

Query Match 74.8%; Score 300; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1; Indels 281;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPVDGPKTKYARDLPRKIDHYRGGQLTPPKAIDFTPGQLLSVSGELG 60
 DB 1 MAKVLCVLYDDPVDGPKTKYARDLPRKIDHYRGGQLTPPKAIDFTPGQLLSVSGELG 60
 QY 61 RPYLESNGHTLVVTSKDDPDSYFERELVDADVISQPPMPAYLTPPERIAKAKNKLALT 120
 DB 61 RPYLESNGHTLVVTSKDDPDSYFERELVDADVISQPPMPAYLTPPERIAKAKNKLALT 120
 QY 121 AGIGSDHVDLQSAIDRNVTVAEYTYCNSISVAEHVVMILSLVRNLPSEHARRKGMNI 180
 DB 121 AGIGSDHVDLQSAIDRNVTVAEYTYCNSISVAEHVVMILSLVRNLPSEHARRKGMNI 180
 QY 181 ADCVSHAUDLEAMHVTGAAGRIGLAVLRRLAFDVLHLYTDRHRLPESVEKELNLTWHA 240
 DB 181 ADCVSHAUDLEAMHVTGAAGRIGLAVLRRLAFDVLHLYTDRHRLPESVEKELNLTWHA 240
 QY 241 TREDMYPCVDVYTLNCLPRLPETEHMINDETLLKFKRGAYIVNTARGKLCDDRAVARALES 300
 DB 241 TREDMYPCVDVYTLNCLPRLPETEHMINDETLLKFKRGAYIVNTARGKLCDDRAVARALES 300
 QY 301 GRLAGYAGDVWFPQAPKDPWMTMPYNGMTPHISGTLTQAARYAAGTREILCEFEGR 360
 DB 301 GRLAGYAGDVWFPQAPKDPWMTMPYNGMTPHISGTLTQAARYAAGTREILCEFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSEBAKFKKAV 401
 DB 361 PIRDEYLIYOGGALAGTGAHYSKGNATGSEBAKFKKAV 401

RESULT 3
 AAU99113
 ID AAU99113 standard; Protein; 401 AA.

XX AAU99113;

DT 24-SEP-2002 (first entry)

DE Mycobacterium vaccae formate dehydrogenase C256S mutant.

XX Formate dehydrogenase; alcohol production; ketone;
 KW mutant; muteln.
 XX
 OS Mycobacterium vaccae.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 256 /note= "Wild type Cys substituted by Ser"
 XX
 PN EPI211316-A1.
 XX
 PD 05-JUN-2002.
 XX
 PP 27-NOV-2001; 2001EP-0128170.
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAICEL) DAICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from Mycobacterium
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX
 PS Claim 8; Page -: 42pp; English.
 XX
 CC This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.
 CC
 XX
 SQ Sequence 401 AA;
 Query Match 74.8%; Score 300; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1.1e-281;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPVGKPYRTYARDLPKIDHYPGQILPTPKAIDFTPGQLGVSSELGL 60
 DB 1 MAKVLCVLYDDPVGKPYRTYARDLPKIDHYPGQILPTPKAIDFTPGQLGVSSELGL 60
 QY 61 RPYLESNGHLLVYTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKMLKALT 120
 DB 61 RPYLESNGHLLVYTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKMLKALT 120
 QY 121 AGISGDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNYLPSHEMARKGWN 180
 DB 121 AGISGDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNYLPSHEMARKGWN 180
 QY 181 ADCVSAHYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYDRRLPESVEKELNLTWHA 240
 DB 181 ADCVSAHYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHYDRRLPESVEKELNLTWHA 240
 QY 241 TREDMYFVCQVVTNLCPRLPTEHMINDETLKLFKRGAYIVNTARGKLCDDAVALAES 300
 DB 241 TREDMYFVCQVVTNLCPRLPTEHMINDETLKLFKRGAYIVNTARGKLCDDAVALAES 300
 QY 301 GRLAGYAGDVWFPQPAKDKHWRMTPTNGMTPHISGTTLTQAQARYAAGTRELCECFEGR 360
 DB 301 GRLAGYAGDVWFPQPAKDKHWRMTPTNGMTPHISGTTLTQAQARYAAGTRELCECFEGR 360
 QY 361 PIRDEYLIYOGGALAGTGAHSHYSKGNATGSESEAKFKKAV 401

DB 361 PIRDEYLIYOGGALAGTGAHSHYSKGNATGSESEAKFKKAV 401
 RESULT 4
 ID AAU99115
 XX AAU99115 standard; Protein; 401 AA.
 AC AAU99115;
 XX
 DT 24-SEP-2002 (first entry)
 DE Mycobacterium vaccae formate dehydrogenase C256V mutant.
 XX
 DE Formate dehydrogenase; alcohol production; ketone;
 KW mutant; muteln.
 XX
 OS Mycobacterium vaccae.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
 XX
 PN EPI211316-A1.
 XX
 PD 05-JUN-2002.
 XX
 PP 27-NOV-2001; 2001EP-0128170.
 XX
 PR 29-NOV-2000; 2000JP-0363894.
 PR 24-AUG-2001; 2001JP-0254631.
 XX
 PA (DAICEL) DAICEL CHEM IND LTD.
 XX
 PI Mitsuhashi K, Yamamoto H, Kimoto N;
 XX
 DR WPI; 2002-464925/50.
 XX
 PT New mutant forms of formate dehydrogenase derived from Mycobacterium
 PT vaccae have enhanced activity in the presence of organic solvents and
 PT are useful for producing alcohols from ketone raw materials
 XX
 PS Claim 8; Page -: 42pp; English.
 XX
 CC This invention relates to a mutant form of the Mycobacterium vaccae
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a Mycobacterium vaccae
 CC formate dehydrogenase protein mutant of the invention.
 CC
 XX
 SQ Sequence 401 AA;
 Query Match 74.8%; Score 300; DB 23; Length 401;
 Best Local Similarity 99.8%; Pred. No. 1.1e-281;
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAKVLCVLYDDPVGKPYRTYARDLPKIDHYPGQILPTPKAIDFTPGQLGVSSELGL 60
 DB 1 MAKVLCVLYDDPVGKPYRTYARDLPKIDHYPGQILPTPKAIDFTPGQLGVSSELGL 60
 QY 61 RPYLESNGHLLVYTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKMLKALT 120
 DB 61 RPYLESNGHLLVYTSDDKGPDSVFERELVDADVVISQPFMPAYLTPERIAKAKMLKALT 120
 QY 121 AGISGDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNYLPSHEMARKGWN 180
 DB 121 AGISGDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNYLPSHEMARKGWN 180

PS Claim 8; Page -: 42pp; English.

XX This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.

XX Sequence 401 AA:

Query Match 73.3%; Score 294; DB 23; Length 401;
Best Local Similarity 99.7%; Pred. No. 6.8e-276;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 VLYDDPYDGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
DB 7 VLYDDPYDGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
QY 67 NGHTLVVTSDDKGDPSVEFERELVDADYVISQFPWPAYLTPERIKAKNKLALTAGIGSD 126
DB 67 NGHTLVVTSDDKGDPSVEFERELVDADYVISQFPWPAYLTPERIKAKNKLALTAGIGSD 126
QY 127 HVDLQSAIDRNVTAAEVTYCNISVAEHVMMILSLVKNYLPSEHMARKGMNTADCVSH 186
DB 127 HVDLQSAIDRNVTAAEVTYCNISVAEHVMMILSLVKNYLPSEHMARKGMNTADCVSH 186
QY 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRRLPESYEKELNTLWHTATREDMY 246
DB 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRRLPESYEKELNTLWHTATREDMY 246
QY 247 PVCDDVVTLCNPLHETEHMINDETLKLFRKRAYIVNTARGKLCRDAAVARALESGRLAGY 306
DB 247 PVCDDVVTLCNPLHETEHMINDETLKLFRKRAYIVNTARGKLCRDAAVARALESGRLAGY 306
QY 307 AGDVWFQAPAKDHPMTMPYNGMTPHISGTTTLTAQAYAACTREILECFEGRPINDEY 366
DB 307 AGDVWFQAPAKDHPMTMPYNGMTPHISGTTTLTAQAYAACTREILECFEGRPINDEY 366
QY 367 LIVOGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401
DB 367 LIVOGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401

RESULT 7
AAU09110 standard; Protein: 401 AA.

XX AAU09110;

XX 24-SEP-2002 (first entry)

XX Mycobacterium vaccae formate dehydrogenase C65/C256A mutant.

XX Formate dehydrogenase; alcohol production; ketone;
XX mutant; mutlein.
XX Mycobacterium vaccae.
XX Synthetic.

XX Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 256 /note= "Wild type Cys substituted by Ala"
XX EPI211316-A1.
XX 05-JUN-2002.

PF 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
PR 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL) DAICEL CHEM IND LTD.
XX Mitsunashi K, Yamamoto H, Kimoto N;
XX WPI; 2002-464925/50.
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials

XX Claim 8; Page -: 42pp; English.

XX This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.

XX Sequence 401 AA:

Query Match 73.3%; Score 294; DB 23; Length 401;
Best Local Similarity 99.7%; Pred. No. 6.8e-276;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 VLYDDPYDGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
DB 7 VLYDDPYDGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGSVSGELGRPYLES 66
QY 67 NGHTLVVTSDDKGDPSVEFERELVDADYVISQFPWPAYLTPERIKAKNKLALTAGIGSD 126
DB 67 NGHTLVVTSDDKGDPSVEFERELVDADYVISQFPWPAYLTPERIKAKNKLALTAGIGSD 126
QY 127 HVDLQSAIDRNVTAAEVTYCNISVAEHVMMILSLVKNYLPSEHMARKGMNTADCVSH 186
DB 127 HVDLQSAIDRNVTAAEVTYCNISVAEHVMMILSLVKNYLPSEHMARKGMNTADCVSH 186
QY 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRRLPESYEKELNTLWHTATREDMY 246
DB 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRRLPESYEKELNTLWHTATREDMY 246
QY 247 PVCDDVVTLCNPLHETEHMINDETLKLFRKRAYIVNTARGKLCRDAAVARALESGRLAGY 306
DB 247 PVCDDVVTLCNPLHETEHMINDETLKLFRKRAYIVNTARGKLCRDAAVARALESGRLAGY 306
QY 307 AGDVWFQAPAKDHPMTMPYNGMTPHISGTTTLTAQAYAACTREILECFEGRPINDEY 366
DB 307 AGDVWFQAPAKDHPMTMPYNGMTPHISGTTTLTAQAYAACTREILECFEGRPINDEY 366
QY 367 LIVOGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401
DB 367 LIVOGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401

RESULT 8
AAU09111 standard; Protein: 401 AA.

XX AAU09111;

XX 24-SEP-2002 (first entry)

XX Mycobacterium vaccae formate dehydrogenase C65/C256V mutant.
XX Formate dehydrogenase; alcohol production; ketone;

```

KW mutant; muteln.
XX Mycobacterium vaccae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild type Cys substituted by Ser"
FT Misc-difference 256
FT /note= "Wild type Cys substituted by Val"
XX
XX EP1211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
XX 24-AUG-2001; 2001JP-0254631.
XX
XX (DAICEL ) DAICEL CHEM IND LTD.
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
XX
XX WPI; 2002-464925/50.
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
XX vaccae have enhanced activity in the presence of organic solvents and
XX are useful for producing alcohols from ketone raw materials
XX
XX Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
XX formate dehydrogenase protein which has strong formate dehydrogenase
XX activity in the presence of an organic solvent. The mutant formate
XX dehydrogenase protein of the invention may be used to produce alcohols
XX using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX proteins, the mutant polypeptides of the invention retain high enzymatic
XX activity in the presence of the alcohol product and so provides a higher
XX product yield. The present sequence represents a Mycobacterium vaccae
XX formate dehydrogenase protein mutant of the invention.
XX
XX Sequence 401 AA:
SQ
Query Match 73.3%; Score 294; DB 23; Length 401;
Best Local Similarity 99.7%; Pred. No. 6,8e-276;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 7 VLYDDPVDDGPKTYARDLPRKIDHYGGQILPPKAIIDFTPGQLLSVSGELGLRPLYLES 66
DB 7 VLYDDPVDDGPKTYARDLPRKIDHYGGQILPPKAIIDFTPGQLLSVSGELGLRPLYLES 66
QY 67 NGHTLVVTSKDGSDPSVFERELVDADVVISQFPWPAYLPERIAKAKNKLALTAGIGSD 126
DB 67 NGHTLVVTSKDGSDPSVFERELVDADVVISQFPWPAYLPERIAKAKNKLALTAGIGSD 126
QY 127 HDLQSAIDRNVTVAEYTCNSISVAEHVYMMILSLVRNLPESHENARKGWNADCVSH 186
DB 127 HDLQSAIDRNVTVAEYTCNSISVAEHVYMMILSLVRNLPESHENARKGWNADCVSH 186
QY 187 AYDLEAMHGTVAAGTIGLAVLRRLARFVNLHYTRHRLPRESEVEKELNTWTATREDMY 246
DB 187 AYDLEAMHGTVAAGTIGLAVLRRLARFVNLHYTRHRLPRESEVEKELNTWTATREDMY 246
QY 247 PUCDDVTTLNCPRLPEREHNINDETLKLFKRGATVMTARGKLCDDAVARALSSGRAGY 306
DB 247 PUCDDVTTLNCPRLPEREHNINDETLKLFKRGATVMTARGKLCDDAVARALSSGRAGY 306
QY 307 AGDWEPQAPAKDHPMRTMPYNGMTPHISGTTLTAQARYAAGTRELIECFEESGRPIRDEY 366
DB 307 AGDWEPQAPAKDHPMRTMPYNGMTPHISGTTLTAQARYAAGTRELIECFEESGRPIRDEY 366
QY 367 LVIYGGALAGTGAHSTSKGNATGSGSEBAKFKKAV 401

```

```

DB 367 LVIYGGALAGTGAHSTSKGNATGSGSEBAKFKKAV 401
|||||
RESULT 9
AAU09117
ID AAU09117 standard; Protein; 401 AA.
XX
XX AAU09117;
XX
XX 24-SEP-2002 (first entry)
XX
XX Mycobacterium vaccae formate dehydrogenase C6a/C256v mutant.
XX
XX Formate dehydrogenase; alcohol production; ketone;
XX mutant; muteln.
XX
XX Mycobacterium vaccae.
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 6 /note= "Wild type Cys substituted by Ala"
FT Misc-difference 256 /note= "Wild type Cys substituted by Val"
XX
XX EP1211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
XX 24-AUG-2001; 2001JP-0254631.
XX
XX (DAICEL ) DAICEL CHEM IND LTD.
XX
XX Mitsuhashi K, Yamamoto H, Kimoto N;
XX
XX WPI; 2002-464925/50.
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
XX vaccae have enhanced activity in the presence of organic solvents and
XX are useful for producing alcohols from ketone raw materials
XX
XX Claim 8; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
XX formate dehydrogenase protein which has strong formate dehydrogenase
XX activity in the presence of an organic solvent. The mutant formate
XX dehydrogenase protein of the invention may be used to produce alcohols
XX using ketones as raw materials. Unlike prior formatted dehydrogenase,
XX proteins, the mutant polypeptides of the invention retain high enzymatic
XX activity in the presence of the alcohol product and so provides a higher
XX product yield. The present sequence represents a Mycobacterium vaccae
XX formate dehydrogenase protein mutant of the invention.
XX
XX Sequence 401 AA:
SQ
Query Match 73.3%; Score 294; DB 23; Length 401;
Best Local Similarity 99.7%; Pred. No. 6,8e-276;
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 7 VLYDDPVDDGPKTYARDLPRKIDHYGGQILPPKAIIDFTPGQLLSVSGELGLRPLYLES 66
DB 7 VLYDDPVDDGPKTYARDLPRKIDHYGGQILPPKAIIDFTPGQLLSVSGELGLRPLYLES 66
QY 67 NGHTLVVTSKDGSDPSVFERELVDADVVISQFPWPAYLPERIAKAKNKLALTAGIGSD 126
DB 67 NGHTLVVTSKDGSDPSVFERELVDADVVISQFPWPAYLPERIAKAKNKLALTAGIGSD 126
QY 127 HDLQSAIDRNVTVAEYTCNSISVAEHVYMMILSLVRNLPESHENARKGWNADCVSH 186

```

```
Db 127 HVDLQSAIDRNVTAEVTYGNSISVAEHVMMILSLVNYLPSHEMARKGMNIADCVSH 186
QY 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESYEKELNTWHTREDMY 246
Db 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESYEKELNTWHTREDMY 246
QY 247 PVCQVVTNLNCPHETETHEIMIDETLKLFRGAYLYVNTARGLCDRDAAVRLAESGRLAGY 306
Db 247 PVCQVVTNLNCPHETETHEIMIDETLKLFRGAYLYVNTARGLCDRDAAVRLAESGRLAGY 306
QY 307 AGDVWFPQAPKDPHWRTPMYPNGMTPHISGTTLTQAARYAAGTREILCEFPFGAPRIRDEY 366
Db 307 AGDVWFPQAPKDPHWRTPMYPNGMTPHISGTTLTQAARYAAGTREILCEFPFGAPRIRDEY 366
QY 367 LIVOGGALACTGAHYSKGNATGSGSEAKFKKAV 401
Db 367 LIVOGGALACTGAHYSKGNATGSGSEAKFKKAV 401

RESULT 10
AAU9121
ID AAU9121 standard; Protein: 401 AA.
AC AAU9121;
XX
XX 24-SEP-2002 (first entry)
DT
XX
DE Mycobacterium vaccae formate dehydrogenase C65/C256V/C355S mutant.
KW Formate dehydrogenase; alcohol production; ketone;
KW mutant; mutlein.
XX
OS Mycobacterium vaccae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 6 /note= "Wild type Cys substituted by Ser"
FT MISC-difference 256 /note= "Wild type Cys substituted by Val"
FT MISC-difference 256 /note= "Wild type Cys substituted by Val"
FT MISC-difference 256 /note= "Wild type Cys substituted by Ser"
PN EPI21J316-A1.
XX
XX 05-JUN-2002.
PD
XX
XX 27-NOV-2001; 2001EP-0128170.
PF
XX
XX 29-NOV-2000; 2000JP-0363894.
PR 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL ) DAICEL CHEM IND LTD.
XX
XX Mitsunashi K, Yamamoto H, Kimoto N;
XX
XX WPI: 2002-464925/50.
DR
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
XX Example 6; Page -: 42pp; English.
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.
```

```
XX
SQ Sequence 401 AA:
Query Match 62.1%; Score 249; DB 23; Length 401;
Best Local Similarity 100.0%; Pred. No. 2,Se-232;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 VLYDDPVVDGYPKYTAARDLPRKIDHYPGQQLIPTPKAIDFTPGQLGSVSGELGRPYLES 66
Db 7 VLYDDPVVDGYPKYTAARDLPRKIDHYPGQQLIPTPKAIDFTPGQLGSVSGELGRPYLES 66
QY 67 NGHTLVYTSDDGDGDSVERELVDADVVISQPFNPAYLTTPRIAKAKLKTALTAGGSD 126
Db 67 NGHTLVYTSDDGDGDSVERELVDADVVISQPFNPAYLTTPRIAKAKLKTALTAGGSD 126
QY 127 HVDLQSAIDRNVTAEVTYGNSISVAEHVMMILSLVNYLPSHEMARKGMNIADCVSH 186
Db 127 HVDLQSAIDRNVTAEVTYGNSISVAEHVMMILSLVNYLPSHEMARKGMNIADCVSH 186
QY 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESYEKELNTWHTREDMY 246
Db 187 AYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTDRHRLPESYEKELNTWHTREDMY 246
QY 247 PVCQVVTNLN 255
Db 247 PVCQVVTNLN 255

RESULT 11
AAU9120
ID AAU9120 standard; Protein: 401 AA.
AC AAU9120;
XX
XX 24-SEP-2002 (first entry)
DT
XX
DE Mycobacterium vaccae formate dehydrogenase C65/C249S/C256V mutant.
KW Formate dehydrogenase; alcohol production; ketone;
KW mutant; mutlein.
XX
OS Mycobacterium vaccae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 6 /note= "Wild type Cys substituted by Ser"
FT MISC-difference 249 /note= "Wild type Cys substituted by Ser"
FT MISC-difference 256 /note= "Wild type Cys substituted by Val"
FT MISC-difference 256 /note= "Wild type Cys substituted by Val"
PN EPI21J316-A1.
XX
XX 05-JUN-2002.
PD
XX
XX 27-NOV-2001; 2001EP-0128170.
PF
XX
XX 29-NOV-2000; 2000JP-0363894.
PR 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL ) DAICEL CHEM IND LTD.
XX
XX Mitsunashi K, Yamamoto H, Kimoto N;
XX
XX WPI: 2002-464925/50.
DR
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
XX Example 5; Page -: 42pp; English.
XX
```

CC This invention relates to a mutant form of the *Mycobacterium vaccae*
 CC formate dehydrogenase protein which has strong formate dehydrogenase
 CC activity in the presence of an organic solvent. The mutant formate
 CC dehydrogenase protein of the invention may be used to produce alcohols
 CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
 CC proteins, the mutant polypeptides of the invention retain high enzymatic
 CC activity in the presence of the alcohol product and so provides a higher
 CC product yield. The present sequence represents a *Mycobacterium vaccae*
 CC formate dehydrogenase protein mutant of the invention.

SO Sequence 401 AA;

Query Match 60.3%; Score 242; DB 23; Length 401;

Best Local Similarity 100.0%; Pred. No. 1.5e-225;

Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 VLVDPPVDGPKRYAADLPLKIDHYGGQILPPKAIIDFPQGLLSVSGELRLPYLES 66
 Db 7 VLVDPPVDGPKRYAADLPLKIDHYGGQILPPKAIIDFPQGLLSVSGELRLPYLES 66
 QY 67 NGHTLVVTSKDGPDSVFERELVDADVISQPFMPAYLTPPERIAKAKNLKLTATAGISD 126
 Db 67 NGHTLVVTSKDGPDSVFERELVDADVISQPFMPAYLTPPERIAKAKNLKLTATAGISD 126
 QY 127 HDVLQSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNYPSEHARKGGMNIADCVSH 186
 Db 127 HDVLQSAIDRNVTVAEYTCNSISVAEHVVMILSLVRNYPSEHARKGGMNIADCVSH 186
 QY 187 AYLLEAHVGTAVAGRIGLAVLRRLAPFDVHLHTYDRHRLPESVEKELNTMTATREDMY 246
 Db 187 AYLLEAHVGTAVAGRIGLAVLRRLAPFDVHLHTYDRHRLPESVEKELNTMTATREDMY 246
 QY 247 PV 248
 Db 247 PV 248

RESULT 12
 AAW37857

ID AAW37857 standard; Protein; 401 AA.

XX AAW37857;

DT 10-AUG-1998 (first entry)

XX Formate dehydrogenase.

XX Formate dehydrogenase; recombinant plasmid; (NAD⁺)-dependent;

KW NADH-dependent.

XX Pseudomonas sp.

OS Pseudomonas sp.

XX Key Location/Qualifiers

FT Misc-difference 31..58 /note- "in this amino acid sequence, residues 49-58

FT Misc-difference 151..178 /note- "appear to have been mis-inserted, and should

FT Misc-difference 151..178 /note- "in this amino acid sequence, residues 169-178

FT Misc-difference 151..178 /note- "appear to have been mis-inserted, and should

XX JP10023896-A.

XX 27-JAN-1998.

XX 19-AUG-1996; 96JP-0217060.

XX 07-MAY-1996; 96JP-0112303.

XX (NIRA) UNITIKA LTD.

XX WPI; 1998-152798/14.

DR N-PSDB; AAV29010.

XX New recombinant plasmid - comprises genes encoding NAD⁺ dependent

PT formate (sic) dehydrogenase and NAD dependent amino acid

PT dehydrogenase

XX Example 1; Pages 9-10; 13pp; Japanese.

XX This is the amino acid sequence of formate dehydrogenase. It is

CC used in the method of the invention to create new recombinant

CC plasmid comprising a gene encoding a NAD⁺ dependent formate

CC dehydrogenase, and a gene encoding an NADH-dependent amino acid

CC dehydrogenase.

SO Sequence 401 AA;

Query Match 55.6%; Score 223; DB 19; Length 401;

Best Local Similarity 100.0%; Pred. No. 3.7e-207;

Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 NINADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTYDRHRLPESVEKELNTM 238
 Db 179 NINADCVSHAYDLEAMHVGTVAAAGRIGLAVLRRLAPFDVHLHTYDRHRLPESVEKELNTM 238
 QY 239 HATREDMPYCDVVTLNCPLHPETEMINDETLEKFKGAYIVNTARGKLCRDAAVARAL 298
 Db 239 HATREDMPYCDVVTLNCPLHPETEMINDETLEKFKGAYIVNTARGKLCRDAAVARAL 298
 QY 299 ESGRLAGYAGDWFPQAPARDHWRMPYNGMTPHISGTTLTQAQYAAAGTRILCCFE 358
 Db 299 ESGRLAGYAGDWFPQAPARDHWRMPYNGMTPHISGTTLTQAQYAAAGTRILCCFE 358
 QY 359 GREIRDEYLIYOGGALAGTAGAHSHSKGNATGSGSEPAKFKKAY 401
 Db 359 GREIRDEYLIYOGGALAGTAGAHSHSKGNATGSGSEPAKFKKAY 401

RESULT 13
 AAU99114

ID AAU99114 standard; Protein; 401 AA.

XX AAU99114;

DT 24-SEP-2002 (first entry)

XX Mycobacterium vaccae formate dehydrogenase C146S/C256S mutant.

XX Formate dehydrogenase; alcohol production; ketone;

KW mutant; mutlein.

XX Mycobacterium vaccae.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 146 /note- "wild type Cys substituted by Ser"

FT Misc-difference 256 /note- "wild type Cys substituted by Ser"

FT Misc-difference 256 /note- "wild type Cys substituted by Ser"

XX EP1211316-A1.

XX 05-JUN-2002.

XX 27-NOV-2001; 2001EP-0128170.

XX 29-NOV-2000; 2000JP-0363894.

XX 24-AUG-2001; 2001JP-0254631.

XX (DAIIC) DAICEL CHEM IND LTD.

XX Mitsuhashi K, Yamamoto H, Kimoto N;

XX WPI; 2002-464925/50.

XX New mutant forms of formate dehydrogenase derived from *Mycobacterium*
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
XX Claim 8; Page -: 42pp; English.
XX
CC This invention relates to a mutant form of the *Mycobacterium vaccae*
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a *Mycobacterium vaccae*
CC formate dehydrogenase protein mutant of the invention.
XX
SQ Sequence 401 AA;
XX
Query Match 49.6%; Score 199; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 6.3e-184;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVDGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSGLGL 60
DB 1 MAKVLCVLYDDPVDGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSGLGL 60
QY 61 RPYLESNGHTLVVTSDDGPDVSFERELVDADVVISOPFMPAYLTPERIAKAKMLKALT 120
DB 61 RPYLESNGHTLVVTSDDGPDVSFERELVDADVVISOPFMPAYLTPERIAKAKMLKALT 120
QY 121 AGISDHDVDSADIRNVTAEVYTCNSISVAEHVMMIISLVNNTLPSEHMARKGWN 180
DB 121 AGISDHDVDSADIRNVTAEVYTCNSISVAEHVMMIISLVNNTLPSEHMARKGWN 180
QY 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPPDVHLHYDRRLPESVEKELNTWHA 240
DB 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPPDVHLHYDRRLPESVEKELNTWHA 240
QY 241 TREDMYPVCVVTNCPLEPTEHINDETLKLRGAYIVNTARGKLCDDAARALES 300
DB 241 TREDMYPVCVVTNCPLEPTEHINDETLKLRGAYIVNTARGKLCDDAARALES 300
QY 301 GRLAGYAGDWFPAPAKDHPWRTMPYNGMTPHISGTTLTQAARYAAGTREILCFFEGR 360
DB 301 GRLAGYAGDWFPAPAKDHPWRTMPYNGMTPHISGTTLTQAARYAAGTREILCFFEGR 360
QY 361 PIRDEYLIYOGALAGTGAHSYSKGNATGSGSEEAKEKAV 401
DB 361 PIRDEYLIYOGALAGTGAHSYSKGNATGSGSEEAKEKAV 401

RESULT 14
AAU99116
ID AAU99116 standard; Protein: 401 AA.
XX
AC AAU99116;
XX
XX 24-SEP-2002 (first entry)
XX
DE *Mycobacterium vaccae* formate dehydrogenase C146S/C256V mutant.
XX
KW Formate dehydrogenase; alcohol production; ketone;
KM mutant; muten.
XX
OS *Mycobacterium vaccae*.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Msc-difference 146 /note- "Wild type Cys substituted by Ser"
FT Msc-difference 256 /note- "Wild type Cys substituted by Val"
FT

XX
PN EPI211316-A1.
XX
XX 05-JUN-2002.
XX
XX 27-NOV-2001; 2001EP-0128170.
XX
XX 29-NOV-2000; 2000JP-0363894.
PR 24-AUG-2001; 2001JP-0254631.
XX
PA (DAI) DAICEL CHEM IND LTD.
XX
PI Mitsuhashi K, Yamamoto H, Kimoto N;
XX
DR WPI; 2002-464925/50.
XX
XX
PT New mutant forms of formate dehydrogenase derived from *Mycobacterium*
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
XX Claim 8; Page -: 42pp; English.
XX
CC This invention relates to a mutant form of the *Mycobacterium vaccae*
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a *Mycobacterium vaccae*
CC formate dehydrogenase protein mutant of the invention.
XX
SQ Sequence 401 AA;
XX
Query Match 49.6%; Score 199; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 6.3e-184;
Matches 399; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAKVLCVLYDDPVDGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSGLGL 60
DB 1 MAKVLCVLYDDPVDGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGQLGVSGLGL 60
QY 61 RPYLESNGHTLVVTSDDGPDVSFERELVDADVVISOPFMPAYLTPERIAKAKMLKALT 120
DB 61 RPYLESNGHTLVVTSDDGPDVSFERELVDADVVISOPFMPAYLTPERIAKAKMLKALT 120
QY 121 AGISDHDVDSADIRNVTAEVYTCNSISVAEHVMMIISLVNNTLPSEHMARKGWN 180
DB 121 AGISDHDVDSADIRNVTAEVYTCNSISVAEHVMMIISLVNNTLPSEHMARKGWN 180
QY 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPPDVHLHYDRRLPESVEKELNTWHA 240
DB 181 ADCVSHAYDLEAMHGVTAAGRIGLAVLRRLAPPDVHLHYDRRLPESVEKELNTWHA 240
QY 241 TREDMYPVCVVTNCPLEPTEHINDETLKLRGAYIVNTARGKLCDDAARALES 300
DB 241 TREDMYPVCVVTNCPLEPTEHINDETLKLRGAYIVNTARGKLCDDAARALES 300
QY 301 GRLAGYAGDWFPAPAKDHPWRTMPYNGMTPHISGTTLTQAARYAAGTREILCFFEGR 360
DB 301 GRLAGYAGDWFPAPAKDHPWRTMPYNGMTPHISGTTLTQAARYAAGTREILCFFEGR 360
QY 361 PIRDEYLIYOGALAGTGAHSYSKGNATGSGSEEAKEKAV 401
DB 361 PIRDEYLIYOGALAGTGAHSYSKGNATGSGSEEAKEKAV 401

RESULT 15
AAU99107
ID AAU99107 standard; Protein: 401 AA.
XX
AC AAU99107;
XX

DT 24-SEP-2002 (first entry)
XX Mycobacterium vaccae formate dehydrogenase C6S/C146S/C256S mutant.
XX
XX Formate dehydrogenase; alcohol production; ketone;
KM mutant; muteln.
XX
XX Mycobacterium vaccae.
OS
XX Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 6 /note= "Wild type Cys substituted by Ser"
FT MISC-difference 146 /note= "Wild type Cys substituted by Ser"
FT MISC-difference 256 /note= "Wild type Cys substituted by Ser"
FT MISC-difference 256 /note= "Wild type Cys substituted by Ser"
XX
XX EPI211316-A1.
PN
XX
XX 05-JUN-2002.
PD
XX
XX 27-NOV-2001; 2001EP-0128170.
PF
XX
XX 29-NOV-2000; 2000JP-0363894.
PR 24-AUG-2001; 2001JP-0254631.
XX
XX (DAIL) DAICEL CHEM IND LTD.
PA
XX
XX Mitsubashi K, Yamamoto H, Kimoto N;
PI
XX WPI; 2002-464925/50.
DR
XX
XX New mutant forms of formate dehydrogenase derived from Mycobacterium
PT vaccae have enhanced activity in the presence of organic solvents and
PT are useful for producing alcohols from ketone raw materials
XX
XX Claim 8; Page -: 42pp; English.
PS
XX
XX This invention relates to a mutant form of the Mycobacterium vaccae
CC formate dehydrogenase protein which has strong formate dehydrogenase
CC activity in the presence of an organic solvent. The mutant formate
CC dehydrogenase protein of the invention may be used to produce alcohols
CC using ketones as raw materials. Unlike prior formatted dehydrogenase,
CC proteins, the mutant polypeptides of the invention retain high enzymatic
CC activity in the presence of the alcohol product and so provides a higher
CC product yield. The present sequence represents a Mycobacterium vaccae
CC formate dehydrogenase protein mutant of the invention.
CC
XX
SQ Sequence 401 AA;

Query Match 48.1%; Score 193; DB 23; Length 401;
Best Local Similarity 99.5%; Pred. No. 4.1e-178;
Matches 393; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 VLYDDPVDTGPKTYARDLFPKIDHYRGGQLPTPKAIDFTPGQLGVSSELGLRPLYLES 66
DB 7 VLYDDPVDTGPKTYARDLFPKIDHYRGGQLPTPKAIDFTPGQLGVSSELGLRPLYLES 66
QY 67 NGHTLVVTSDDKSDPSVEFERELVDADVVISQPFMPAYLTPERTAKAKNLKALTAGIGSD 126
DB 67 NGHTLVVTSDDKSDPSVEFERELVDADVVISQPFMPAYLTPERTAKAKNLKALTAGIGSD 126
QY 127 HVDLQSAIDRNVAEVTYCNISVAEHVYMMILSLVRNLPSEHNAKRGWNIADCVSH 186
DB 127 HVDLQSAIDRNVAEVTYCNISVAEHVYMMILSLVRNLPSEHNAKRGWNIADCVSH 186
QY 187 AYDLAMHVGTVAAAGTGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTWHATREDMY 246
DB 187 AYDLAMHVGTVAAAGTGLAVLRRLAPFDVHLHYTDRHRLPESVEKELNTWHATREDMY 246
QY 247 PVEDVVTLNCPLPETEHMINDETLLKFRGATVVTARGKLCDDRAVARALESGRLAGY 306
DB 247 PVEDVVTLNCPLPETEHMINDETLLKFRGATVVTARGKLCDDRAVARALESGRLAGY 306

DB 247 PVEDVVTLNCPLPETEHMINDETLLKFRGATVVTARGKLCDDRAVARALESGRLAGY 306
QY 307 AGDVWFQAPAKDHPMTMTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGRPIINDEX 366
DB 307 AGDVWFQAPAKDHPMTMTMPYNGMTPHISGTTLTAQARYAAGTREILECFEGRPIINDEX 366
QY 367 LIVOGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401
DB 367 LIVOGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401

Search completed: July 24, 2003, 20:59:04
Job time : 85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:44:28 ; Search time 30 Seconds

(Without alignments)
565.555 Million cell updates/sec

Title: US-09-996-008b-2

Perfect score: 401

Sequence: 1 MAKVLCVLYDDPVDGYPKRY.....YSGNATGSGSEAKFKKAV 401

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCrus.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	4.0	364	US-08-817-926-51	Sequence 51, Appl
2	16	4.0	364	US-09-203-893A-2	Sequence 2, Appl
3	16	4.0	364	US-09-203-893A-4	Sequence 4, Appl
4	16	4.0	364	US-09-203-893A-6	Sequence 6, Appl
5	16	4.0	364	US-09-203-893A-8	Sequence 8, Appl
6	16	4.0	364	US-09-203-893A-10	Sequence 10, Appl
7	16	4.0	364	US-09-203-893A-12	Sequence 12, Appl
8	16	4.0	364	US-09-203-893A-14	Sequence 14, Appl
9	16	4.0	364	US-09-203-893A-16	Sequence 16, Appl
10	16	4.0	364	US-09-203-893A-18	Sequence 18, Appl
11	16	4.0	364	US-09-203-893A-20	Sequence 20, Appl
12	16	4.0	364	US-09-203-893A-22	Sequence 22, Appl
13	16	4.0	364	US-09-203-893A-24	Sequence 24, Appl
14	16	4.0	364	US-09-203-893A-26	Sequence 26, Appl
15	16	4.0	364	US-09-203-893A-28	Sequence 28, Appl
16	16	4.0	364	US-09-203-893A-30	Sequence 30, Appl
17	16	4.0	364	US-09-203-893A-32	Sequence 32, Appl
18	8	2.0	253	US-09-530-058-6	Sequence 6, Appl
19	8	2.0	274	US-09-252-991A-24164	Sequence 2, Appl
20	8	2.0	550	US-08-844-058-2	Sequence 2, Appl
21	7	1.7	20	US-08-290-736C-16	Sequence 16, Appl
22	7	1.7	20	US-08-290-736C-27	Sequence 27, Appl
23	7	1.7	133	US-09-384-162-16	Sequence 16, Appl
24	7	1.7	155	US-09-252-991A-17465	Sequence 17465, A
25	7	1.7	157	US-09-252-991A-26320	Sequence 26320, A
26	7	1.7	176	US-09-462-842-3	Sequence 3, Appl
27	7	1.7	182	US-09-040-229B-4	Sequence 4, Appl

28	7	1.7	245	US-09-205-258-369	Sequence 369, App
29	7	1.7	259	US-09-252-991A-26709	Sequence 26709, A
30	7	1.7	268	US-08-965-056-71	Sequence 71, Appl
31	7	1.7	274	US-09-328-352-5762	Sequence 5762, Ap
32	7	1.7	304	US-09-634-238-238	Sequence 238, App
33	7	1.7	316	US-08-414-926A-22	Sequence 22, Appl
34	7	1.7	316	US-08-926-922-22	Sequence 22, Appl
35	7	1.7	316	US-09-253-682-22	Sequence 22, Appl
36	7	1.7	316	US-09-527-657-22	Sequence 22, Appl
37	7	1.7	331	US-09-252-991A-24420	Sequence 24420, A
38	7	1.7	348	US-09-107-532A-5765	Sequence 5765, Ap
39	7	1.7	365	US-09-252-991A-31439	Sequence 31439, A
40	7	1.7	366	US-09-252-991A-32332	Sequence 32332, A
41	7	1.7	394	US-08-867-030B-11	Sequence 11, Appl
42	7	1.7	394	PCT-US95-06119-11	Sequence 11, Appl
43	7	1.7	406	US-09-252-991A-19182	Sequence 19182, A
44	7	1.7	415	US-09-252-991A-19206	Sequence 19206, A
45	7	1.7	424	US-09-328-352-4427	Sequence 4427, Ap

ALIGNMENTS

RESULT 1
US-08-817-926-51
Sequence 51, Application US/08817926
Patent No. 6001590
GENERAL INFORMATION:
APPLICANT: Komeda, Yoshihiro
APPLICANT: Suda, Hisako
APPLICANT: Tamai, Yukio
APPLICANT: Iwamatsu, Akihito
APPLICANT: Kato, No. 6001590uo
APPLICANT: Sakai, Yasuyoshi
TITLE OF INVENTION: PROMOTER/TERMINATOR FOR CANDIDA BOIDINITII
TITLE OF INVENTION: FORMATE DEHYDROGENASE GENE
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817, 926
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/02597
FILING DATE: 12-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 234133/1995
FILING DATE: 12-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 42536/1996
FILING DATE: 29-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 081356/0112
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 364 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-926-51

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 2
US-09-203-893A-2
Sequence 2, Application US/09203893A
Patent No. 6242234
GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of
TITLE OF INVENTION: the new Formate dehydrogenases
FILE REFERENCE: DEGU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203.893A
CURRENT FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
PRIOR FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 364
TYPE: PRT
ORGANISM: Candida boidin1
US-09-203-893A-2

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 3
US-09-203-893A-4
Sequence 4, Application US/09203893A
Patent No. 6242234
GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of
TITLE OF INVENTION: the new Formate dehydrogenases
FILE REFERENCE: DEGU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203.893A
CURRENT FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
PRIOR FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 364
TYPE: PRT
ORGANISM: Candida boidin1
US-09-203-893A-4

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

DB 281 GDVWFPQAPKDPHWR 296

RESULT 4
US-09-203-893A-6
Sequence 6, Application US/09203893A
Patent No. 6242234
GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of
TITLE OF INVENTION: the new Formate dehydrogenases
FILE REFERENCE: DEGU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203.893A
CURRENT FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
PRIOR FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 364
TYPE: PRT
ORGANISM: Candida boidin1
US-09-203-893A-6

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 5
US-09-203-893A-8
Sequence 8, Application US/09203893A
Patent No. 6242234
GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of
TITLE OF INVENTION: the new Formate dehydrogenases
FILE REFERENCE: DEGU 3.0-012
CURRENT APPLICATION NUMBER: US/09/203.893A
CURRENT FILING DATE: 1998-12-02
PRIOR APPLICATION NUMBER: DE 19753350.7
PRIOR FILING DATE: 1997-12-03
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 364
TYPE: PRT
ORGANISM: Candida boidin1
US-09-203-893A-8

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 6
US-09-203-893A-10
Sequence 10, Application US/09203893A
Patent No. 6242234
GENERAL INFORMATION:
APPLICANT: Kula, Maria-Regina
TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
TITLE OF INVENTION: boidin1, new Gene sequences encoding these and use of


```

; TITLE OF INVENTION: the new Formate dehydrogenases
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidin11
US-09-203-893A-10

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPPOPAPKDHPR 323
DB 281 GDVWFPPOPAPKDHPR 296

RESULT 7
US-09-203-893A-12
; Sequence 12, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin11, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidin11
US-09-203-893A-12

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPPOPAPKDHPR 323
DB 281 GDVWFPPOPAPKDHPR 296

RESULT 8
US-09-203-893A-14
; Sequence 14, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin11, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 364
; TYPE: PRT
```

```

; ORGANISM: Candida boidin11
US-09-203-893A-14

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPPOPAPKDHPR 323
DB 281 GDVWFPPOPAPKDHPR 296

RESULT 9
US-09-203-893A-16
; Sequence 16, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin11, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidin11
US-09-203-893A-16

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPPOPAPKDHPR 323
DB 281 GDVWFPPOPAPKDHPR 296

RESULT 10
US-09-203-893A-18
; Sequence 18, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boidin11, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boidin11
US-09-203-893A-18

Query Match
Best Local Similarity 100.0%; Score 16; DB 3; Length 364;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPPOPAPKDHPR 323
DB 281 GDVWFPPOPAPKDHPR 296
```

```
RESULT 11
US-09-203-893A-20
; Sequence 20, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 20
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-20
```

```
Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 308 GDVFPQAPKDHPR 323
Db 281 GDVFPQAPKDHPR 296
```

```
RESULT 12
US-09-203-893A-22
; Sequence 22, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 22
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-22
```

```
Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 308 GDVFPQAPKDHPR 323
Db 281 GDVFPQAPKDHPR 296
```

```
RESULT 13
US-09-203-893A-24
; Sequence 24, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
```

```
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 24
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-24
```

```
Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 308 GDVFPQAPKDHPR 323
Db 281 GDVFPQAPKDHPR 296
```

```
RESULT 14
US-09-203-893A-26
; Sequence 26, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 26
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-26
```

```
Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 308 GDVFPQAPKDHPR 323
Db 281 GDVFPQAPKDHPR 296
```

```
RESULT 15
US-09-203-893A-28
; Sequence 28, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 28
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-28
```

Query Match 4.0%; Score 16; DB 3; Length 364;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 308 GDVWFPOPAPKDHPR 323
 ||||||||||||
 Db 281 GDVWFPOPAPKDHPR 296

Search completed: July 24, 2003, 20:49:56
 Job time : 42 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:48:38 ; Search time 54 Seconds

(without alignments)
881.901 Million cell updates/sec

Title: US-09-996-008B-2

Perfect score: 401

Sequence: 1 MAKVLCVLYDDPVVDGYPKTY.....YSGNATGSGSEAKFKKAV 401

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Word size : 0

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA.*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCr_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCrUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	ID	Description
1	19	4.7	387	15	US-10-156-761-9479	Sequence 9479, Ap	
2	11	2.7	376	15	US-10-224-567-1	Sequence 1, Appl1	
3	2	2.2	333	15	US-10-156-761-12108	Sequence 12108, A	
4	8	2.0	335	9	US-09-815-242-5830	Sequence 5830, Ap	
5	8	2.0	343	9	US-09-815-242-12981	Sequence 12981, A	
6	8	2.0	343	9	US-09-815-242-13150	Sequence 13150, A	
7	8	2.0	398	10	US-09-712-363-160	Sequence 160, App	
8	8	2.0	617	9	US-09-815-242-13228	Sequence 13228, A	
9	8	2.0	764	15	US-10-066-551-2	Sequence 2, Appl1	
10	8	2.0	1176	15	US-10-156-761-11086	Sequence 11086, A	
11	7	1.7	38	15	US-10-099-766-3	Sequence 3, Appl1	
12	7	1.7	40	9	US-09-682-706-6	Sequence 6, Appl1	
13	7	1.7	161	9	US-09-682-706-5	Sequence 5, Appl1	
14	7	1.7	161	15	US-10-099-766-2	Sequence 2, Appl1	
15	7	1.7	162	15	US-10-156-761-13701	Sequence 13701, A	

16	7	1.7	175	15	US-10-156-761-14883	Sequence 14883, A
17	7	1.7	176	15	US-10-339-278-3	Sequence 3, Appl1
18	7	1.7	182	14	US-10-159-749-4	Sequence 4, Appl1
19	7	1.7	218	15	US-10-102-806-515	Sequence 515, App
20	7	1.7	220	10	US-09-947-316-2	Sequence 2, Appl1
21	7	1.7	231	11	US-09-934-455-286	Sequence 286, App
22	7	1.7	245	9	US-09-799-777-57	Sequence 57, Appl1
23	7	1.7	245	10	US-09-947-316-1	Sequence 1, Appl1
24	7	1.7	245	15	US-10-022-282-369	Sequence 369, App
25	7	1.7	258	15	US-10-136-761-7639	Sequence 7639, App
26	7	1.7	268	10	US-09-854-816-71	Sequence 71, Appl1
27	7	1.7	268	15	US-10-156-761-15022	Sequence 15022, A
28	7	1.7	301	10	US-09-971-536-51	Sequence 51, Appl1
29	7	1.7	363	15	US-10-156-761-13930	Sequence 13930, A
30	7	1.7	478	15	US-10-156-761-13018	Sequence 13018, A
31	7	1.7	511	9	US-09-815-242-5147	Sequence 5147, Ap
32	7	1.7	513	9	US-09-833-745-64	Sequence 64, Appl1
33	7	1.7	637	9	US-09-817-310-2	Sequence 2, Appl1
34	7	1.7	658	9	US-09-815-242-10947	Sequence 10947, A
35	7	1.7	663	15	US-10-156-761-15035	Sequence 15035, A
36	7	1.7	679	10	US-09-728-626-5995	Sequence 5995, App
37	7	1.7	692	10	US-10-136-761-14649	Sequence 14649, A
38	7	1.7	731	15	US-10-205-823-502	Sequence 202, App
39	7	1.7	758	10	US-09-735-101-2	Sequence 5, Appl1
40	7	1.7	1203	9	US-09-799-875-5	Sequence 2, Appl1
41	7	1.7	1257	15	US-10-109-324-2	Sequence 3707, Ap
42	7	1.7	1510	10	US-09-728-626-3707	Sequence 3707, Ap
43	7	1.7	1516	15	US-10-136-761-13716	Sequence 13716, A
44	7	1.7	1724	10	US-09-964-899-43	Sequence 43, Appl1
45	7	1.7	1925	15	US-10-205-032-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-10-156-761-9479
; Sequence 9479, Application US/10156761
; Publication NO. US20030119018A1
GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9479
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9479
Query Match 4.7%; Score 19; DB 15; Length 387;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 IDFTPGQLGSGVSGELGR 61
|||||
Db 43 IDFTPGQLGSGVSGELGR 61
|||||
RESULT 2
US-10-224-567-1

```
Sequence 1, Application US/10224567
Publication No. US20030087763A1
GENERAL INFORMATION:
APPLICANT: Kobayashi, Akio
APPLICANT: Fukusaki, Eiichiro
APPLICANT: Isogai, Akira
TITLE OF INVENTION: Method to Promote Growth of a Plant
FILE REFERENCE: 026350-077
CURRENT APPLICATION NUMBER: US/10/224,567
CURRENT FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: US 09/517,427
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: JP 11-56,776
PRIOR FILING DATE: 1999-03-04
PRIOR APPLICATION NUMBER: JP 2000-36,153
PRIOR FILING DATE: 2000-02-15
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 376
TYPE: PRT
ORGANISM: Rice formate dehydrogenase
FEATURE:
NAME/KEY: VARIANT
LOCATION: 287
OTHER INFORMATION: Xaa = Any Amino Acid
US-10-224-567-1
```

```
Query Match 2.7%: Score 11; DB 15; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 308 GDVWFPQAPK 318
Db 305 GDVWFPQAPK 315
```

```
RESULT 3
US-10-156-761-12108
Sequence 12108, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMIURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JIN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTOH, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 12108
LENGTH: 333
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-12108
```

```
Query Match 2.2%: Score 9; DB 15; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 290 DRDAVARAL 298
Db 42 DRDAVARAL 50
```

```
RESULT 4
US-09-815-242-5830
Sequence 5830, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5830
LENGTH: 335
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5830
```

```
Query Match 2.0%: Score 8; DB 9; Length 335;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 319 DHPWRMP 326
Db 277 DHPWRMP 284
```

```
RESULT 5
US-09-815-242-12981
Sequence 12981, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
```

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12981
LENGTH: 343
TYPE: PRT
ORGANISM: *Staphylococcus aureus*
US-09-815-242-12981

Query Match
Best Local Similarity 100.0%; Score 8; DB 9; Length 343;
Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 DHPWRTMP 326
DB 280 DHPWRTMP 287

RESULT 6
US-09-815-242-13150
Sequence 13150, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13150
LENGTH: 343
TYPE: PRT
ORGANISM: *Staphylococcus aureus*
US-09-815-242-13150

Query Match
Best Local Similarity 100.0%; Score 8; DB 9; Length 343;
Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 DHPWRTMP 326
DB 280 DHPWRTMP 287

RESULT 7
US-09-712-363-160
Sequence 160, Application US/09712363
Patent No. US20020164588A1
GENERAL INFORMATION:
APPLICANT: Eisenberg, David
APPLICANT: Rothenstein, Sergio H.
APPLICANT: Marcotte, Edward M.
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
FILE REFERENCE: 07419-033001
CURRENT FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: 60/117,844
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/118,206
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: 60/126,593
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/134,093
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/134,092
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/165,124
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/165,086
PRIOR FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 292
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 160
LENGTH: 398
TYPE: PRT
ORGANISM: *Mycobacterium tuberculosis*
US-09-712-363-160

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 398;
Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 AVLRLAP 213
DB 181 AVLRLAP 188

RESULT 8
US-09-815-242-13228
Sequence 13228, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13228
; LENGTH: 617
; TYPE: PRF
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13228
```

```
Query Match          2.0%; Score 8; DB 9; Length 617;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      211 LAPDVHL 218
        |||||
Db      518 LAPDVHL 525
```

```
RESULT 9
US-10-066-551-2
; Sequence 2, Application US/10066551
; Publication No. US20030100071A1
; GENERAL INFORMATION:
; APPLICANT: Apicella, M. A.
; APPLICANT: Edwards, J. L.
; APPLICANT: Gibson, B. W.
; APPLICANT: Scheffler, K.
; APPLICANT: Brown, E.
; TITLE OF INVENTION: Vaccine and compositions for the
; TITLE OF INVENTION: Prevention and treatment of Neisserial infections
; FILE REFERENCE: 875.045051
; CURRENT APPLICATION NUMBER: US/10/066,551
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/344,452
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/310,356
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/266,070
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 12
; SEQ ID NO 2
; LENGTH: 764
; TYPE: PRF
; ORGANISM: Neisseria gonorrhoeae
US-10-066-551-2
```

```
Query Match          2.0%; Score 8; DB 15; Length 764;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      54 VSGELGR 61
        |||||
Db      178 VSGELGR 185
```

```
RESULT 10
US-10-156-761-11086
; Sequence 11086, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
```

```
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11086
; LENGTH: 1176
; TYPE: PRF
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11086
```

```
Query Match          2.0%; Score 8; DB 15; Length 1176;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      205 LAVLRRLA 212
        |||||
Db      106 LAVLRRLA 113
```

```
RESULT 11
US-10-099-766-3
; Sequence 3, Application US/10099766
; Publication No. US20030036507A1
; GENERAL INFORMATION:
; APPLICANT: Lewis, Kathy
; APPLICANT: Vale, Wylie
; APPLICANT: Marilyn H. Perrin
; APPLICANT: Jean E. Rivier
; APPLICANT: Kotchi S. Kunitake
; APPLICANT: Joszet Gulyas
; TITLE OF INVENTION: Urocortin III and Uses Thereof
; FILE REFERENCE: D6390
; CURRENT APPLICATION NUMBER: US/10/099,766
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/276,069
; PRIOR FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 3
; LENGTH: 38
; TYPE: PRF
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Human urocortin III (hucn III)
US-10-099-766-3
```

```
Query Match          1.7%; Score 7; DB 15; Length 38;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      109 IAKAKNL 115
        |||||
Db      18 IAKAKNL 24
```

```
RESULT 12
US-09-682-706-6
; Sequence 6, Application US/09682706
; Patent No. US20020082409A1
; GENERAL INFORMATION:
; APPLICANT: Hsu, Sheau-yu
; APPLICANT: Hsueh, Aaron
; TITLE OF INVENTION: Stresscopins and their ses
; FILE REFERENCE: STAN210
; CURRENT APPLICATION NUMBER: US/09/682,706
; CURRENT FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/276,615
; PRIOR FILING DATE: 2001-03-15
```



```

; PRIOR APPLICATION NUMBER: 60/244,128
; PRIOR FILING DATE: 2000-10-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 6
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-682-706-6

Query Match
Best Local Similarity 1.7%; Score 7; DB 9; Length 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 IAKAKNL 115
Db 20 IAKAKNL 26

RESULT 13
US-09-682-706-5
; Sequence 5, Application US/09682706
; Patent No. US20020082409A1
; GENERAL INFORMATION:
; APPLICANT: Hsu, Sheau-Yu
; APPLICANT: Hsueh, Aaron
; TITLE OF INVENTION: Stresscoping and their ses
; FILE REFERENCE: STRAN210
; CURRENT APPLICATION NUMBER: US/09/682,706
; CURRENT FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/276,615
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/244,128
; PRIOR FILING DATE: 2000-10-26
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-682-706-5

Query Match
Best Local Similarity 1.7%; Score 7; DB 9; Length 161;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 IAKAKNL 115
Db 137 IAKAKNL 143

RESULT 14
US-10-099-766-2
; Sequence 2, Application US/10099766
; Publication No. US20030036507A1
; GENERAL INFORMATION:
; APPLICANT: Lewis, Kathy
; APPLICANT: Vale, Wylie
; APPLICANT: Marilyn H. Perrin
; APPLICANT: Jean E. Rivier
; APPLICANT: Koichi S. Kunitake
; APPLICANT: Jozsef Guylas
; TITLE OF INVENTION: Urocortin III and Uses Thereof
; FILE REFERENCE: D6390
; CURRENT APPLICATION NUMBER: US/10/099,766
; CURRENT FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/276,069
; PRIOR FILING DATE: 2001-03-15
; NUMBER OF SEQ ID NOS: 17
; SEQ ID NO 2
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```

; FEATURE:
; OTHER INFORMATION: Human urocortin III precursor
US-10-099-766-2

Query Match
Best Local Similarity 1.7%; Score 7; DB 15; Length 161;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 IAKAKNL 115
Db 137 IAKAKNL 143

RESULT 15
US-10-156-761-13701
; Sequence 13701, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272657
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13701
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13701

Query Match
Best Local Similarity 1.7%; Score 7; DB 15; Length 162;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 AVLRLA 212
Db 32 AVLRLA 38
```

Search completed: July 24, 2003, 20:57:39
Job time : 62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: July 24, 2003, 20:56:44 ; Search time 41 seconds
(without alignments)
940.577 Million cell updates/sec

Title: US-09-996-008B-2

Perfect score: 401

Sequence: 1 MAKVLCVLYDDPVDGPKTY.....YSKGNATGGSEAAKFKKAV 401

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched.: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR_76:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	110	27.4	393	2 JU0334	formate dehydrogen
2	24	6.0	401	2 C95293	probable NAD-depen
3	23	5.7	400	2 J7815	formate dehydrogen
4	16	4.0	364	2 J4252	formate dehydrogen
5	12	3.0	236	2 S65308	hypothetical prote
6	12	3.0	376	2 S67300	probable membrane
7	12	3.0	379	2 J02272	formate dehydrogen
8	11	2.7	375	2 A47117	formate dehydrogen
9	11	2.7	377	2 S30088	actA protein - Eme
10	11	2.7	384	2 T31423	formate dehydrogen
11	8	2.0	89	2 D90271	conserved hypotet
12	8	2.0	334	2 AH2585	hypothetical prote
13	8	2.0	334	2 G97367	probable dehydroge
14	8	2.0	374	2 E89779	NAD-dependent form
15	8	2.0	398	2 B70752	hypothetical prote
16	8	2.0	617	2 A95031	prolyl-trRNA synthe
17	8	2.0	617	2 C97902	proline-trRNA ligas
18	8	2.0	764	2 D81883	probable outer-mem
19	8	2.0	1105	2 C87622	TonB-dependent rec
20	7	1.7	92	2 T03677	plt2 protein (clon
21	7	1.7	94	2 H70654	hypothetical prote
22	7	1.7	126	2 D90221	hypothetical prote
23	7	1.7	134	2 E64375	hypothetical prote
24	7	1.7	136	2 D96742	unknown protein.F1
25	7	1.7	138	2 S35233	CAP3-1 protein - S
26	7	1.7	145	2 S55309	probable membrane
27	7	1.7	146	2 D82748	stringent starvati
28	7	1.7	173	2 F72302	conserved hypotet
29	7	1.7	176	1 F64864	protein-disulfide

30	7	1.7	176	2 H90838	protein-disulfide
31	7	1.7	176	2 H85696	hypothetical prote
32	7	1.7	182	2 AC1653	Listeria prophage
33	7	1.7	187	2 B83695	hypothetical prote
34	7	1.7	190	2 G86817	hypothetical prote
35	7	1.7	193	2 G83901	hypothetical prote
36	7	1.7	195	2 E82857	pilX protein XF003
37	7	1.7	208	1 B69066	conserved hypotet
38	7	1.7	210	2 A64161	hypothetical prote
39	7	1.7	212	1 RSHSLH	ribosomal protein
40	7	1.7	212	2 H84266	50S ribosomal prot
41	7	1.7	231	2 T45579	hypothetical prote
42	7	1.7	232	2 T04354	hypothetical prote
43	7	1.7	251	2 C83083	conserved hypotet
44	7	1.7	259	2 T34637	probable type IV p
45	7	1.7	277	2 AE3351	glutamate racemase

ALIGNMENTS

RESULT 1

JU0334

formate dehydrogenase (EC 1.2.1.2) - Pseudomonas sp.

N:Alternate names: NAD-dependent formate dehydrogenase

C:Species: Pseudomonas sp.

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 23-Sep-2002

C:Accession: JU0334

R:Popov, V.O.; Shumilin, I.A.; Ustinnikova, T.B.; Lamzin, V.S.; Egorov, T.A.

Bioorg. Khim. 16, 324-335, 1990

A:Title: NAD-dependent formate dehydrogenase from methylotrophic bacterium Pseudomona

A:Reference number: JU0334; MUID:90290536; PMID:2357236

A:Accession: JU0334

A:Molecule type: protein

A:Residues: 1-393 <POP>

A:Note: article in Russian with English abstract

C:Superfamily: Neurospora formate dehydrogenase

C:Keywords: homodimer; NAD; oxidoreductase

Query Match 27.4%; Score 110; DB 2; Length 393;

Best Local Similarity 100.0%; Pred. No. 5.1e-103;

Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 218 LHYDRHRLPESVEKELNLTWHATREDMYPVCDVVTLNCPHPETEHINDETLLKFRG 277

Db 217 LHYDRHRLPESVEKELNLTWHATREDMYPVCDVVTLNCPHPETEHINDETLLKFRG 276

Qy 278 AYIVNTARGKLCRDVAVARALESGRLAGYAGDVWFPQAPKDPHWPRTMPY 327

Db 277 AYIVNTARGKLCRDVAVARALESGRLAGYAGDVWFPQAPKDPHWPRTMPY 326

RESULT 2

C95293

probable NAD-dependent formate dehydrogenase [imported] - Sinorhizobium meliloti (str

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 23-Sep-2002

C:Accession: C95293

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B

; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meli

A:Reference number: A55262; MUID:21396509; PMID:11481432

A:Accession: C95293

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-401 <KUR>

A:Cross-references: GB:AE006469; PIDN:AAK64909.1; PID:g14523329; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSynA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebaule, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: Sma0478
A;Genome: plasmid
C;Superfamily: Neurospora formate dehydrogenase

Query Match 6.0%; Score 24; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 7.2e-16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 GHTLVVTSKDGSDSVFERELVDA 91
|||||
Db 70 GHTLVVTSKDGSDSVFERELVDA 93

RESULT 3

JC7815
formate dehydrogenase (EC 1.2.1.2) - *Paracoccus* sp. (Strain 12-A)
C;Species: *Paracoccus* sp.
C;Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
C;Accession: JC7815
R;Shinoda, T.; Satoh, T.; Mineki, S.; Iida, M.; Taguchi, H.
BioSci. Biotechnol. Biochem. 66, 271-276, 2002
A;Title: Cloning, nucleotide sequencing, and expression in *Escherichia coli* of the gene
A;Reference number: JC7815; PMID:1199398; MUID:21994041
A;Accession: JC7815
A;Molecule type: DNA
A;Residues: 1-400 <SHI>
A;Cross-references: DDBJ:AB071373
C;Comment: This enzyme, which catalyzes the conversion of formate to carbon dioxide with
ti-enzyme systems, such as bioreactors involving NADH as a coenzyme, and plays a key rol
C;Genetics:
A;Gene: fdh
C;Keywords: oxidoreductase

Query Match 5.7%; Score 23; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 GAHSYKGNATGGSEEAFFKA 400
|||||
Db 378 GAHSYKGNATGGSEEAFFKA 400

RESULT 4

JC4252
formate dehydrogenase (EC 1.2.1.2) - yeast (*Candida methylca*)
C;Species: *Candida methylca*
C;Date: 19-Oct-1995 #sequence_revision 08-Feb-1996 #text_change 23-Sep-2002
C;Accession: JC4252; PC4070
R;Allen, S.J.; Holbrook, J.J.
Gene 162, 99-104, 1995
A;Title: Isolation, sequence and overexpression of the gene encoding NAD-dependent forma
A;Reference number: JC4252; MUID:96009885; PMID:7557425
A;Accession: JC4252
A;Molecule type: DNA
A;Residues: 1-364 <ALL>
A;Cross-references: EMBL:X81129; MUID:g1181203; PIDN:CAA57036.1; PID:g1181204
A;Accession: PC4070
A;Molecule type: protein
A;Residues: 1-30 <AL2>
A;Experimental source: ATCC 56294
C;Comment: This enzyme catalyses the final step in the dissimilatory pathway of methanol
of NAD+ to NADH.
C;Genetics:
A;Gene: fdh
C;Superfamily: Neurospora formate dehydrogenase
C;Keywords: homodimer; NAD; oxidoreductase
F;165-195/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 4.0%; Score 16; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 8.4e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDWFFQAPKDPKDPWR 323
|||||
Db 281 GDWFFQAPKDPKDPWR 296

RESULT 5

S65308
hypothetical protein YPL275w - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein P0326
C;Species: *Saccharomyces cerevisiae*
C;Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 23-Sep-2002
A;Accession: S65308; S65329
R;Duesterhoef, A.; Floeth, M.; Frit, M.; Hilbert, H.; Moestl, D.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S65292
A;Accession: S65308
A;Molecule type: DNA
A;Residues: 1-236 <DUE>
A;Cross-references: EMBL:273632; MUID:g1370566; PIDN:CAA98013.1; PID:g1370568; MIPS:YP
A;Experimental source: strain S288C (AB972)
R;Dellius, H.; Hebling, U.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64967
A;Accession: S65329
A;Molecule type: DNA
A;Residues: 1-236
A;Cross-references: EMBL:273632; MUID:g1370566; PIDN:CAA98013.1; PID:g1370568; MIPS:YP
A;Experimental source: strain S288C (AB972)
C;Genetics:
A;Cross-references: SGD:S0006196
A;Map position: 16L
C;Superfamily: Neurospora formate dehydrogenase

Query Match 3.0%; Score 12; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 QPAPKDPKDPKDPWR 325
|||||
Db 161 QPAPKDPKDPKDPWR 172

RESULT 6

S67300
probable membrane protein YOR388c - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein O6778
C;Species: *Saccharomyces cerevisiae*
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 23-Sep-2002
C;Accession: S67300
R;Dellius, H.; Hebling, U.; Hofmann, B.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67261
A;Accession: S67300
A;Molecule type: DNA
A;Residues: 1-376
A;Cross-references: EMBL:275296; MUID:g1420834; PIDN:CAA99720.1; PID:g1420835; MIPS:YO
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:FDH1
A;Cross-references: SGD:S0005915; MIPS:YOR388c
A;Map position: 15R
C;Superfamily: Neurospora formate dehydrogenase
C;Keywords: transmembrane protein
F;121-137/Domain: transmembrane #status predicted <TM>

Query Match 3.0%; Score 12; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.00098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 314 QPAPKDPHPRM 325
DB 301 QPAPKDPHPRM 312

RESULT 7

formate dehydrogenase (EC 1.2.1.2) precursor, mitochondrial - potato
C:Species: Solanum tuberosum (potato)
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 23-Sep-2002
C:Accession: JQ2272; PC2189; S3125
R:des France-Small, C.C.; Ambard-Bretteville, F.; Small, I.D.; Remy, R.
Plant Physiol. 102, 1171-1177, 1993
A:Title: Identification of a major soluble protein in mitochondria from nonphotosynthetic
A:Reference number: JQ2272; MUID:94105343; PMID:8278546
A:Accession: JQ2272
A:Molecule type: mRNA
A:Residues: 1-379 <DESI>
A:Cross-references: EMBL:Z21493; NID:g297797; PIDN:CAA79702.1; PID:g297798
A:Accession: PC2189
A:Molecule type: protein
A:Residues: 24-50 <DE52>
C:Superfamily: Neurospora formate dehydrogenase
C:Keywords: mitochondrion; oxidoreductase
F:1-23/Domain: transit peptide (mitochondrion) #status predicted <TM>
F:24-379/Product: formate dehydrogenase #status predicted <MAT>
F:199-222/Region: NAD binding #status predicted

Query Match

Best Local Similarity 3.0%; Score 12; DB 2; Length 379;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 119 LTAGISDHYDL 130
DB 119 LTAGISDHYDL 130

RESULT 8

formate dehydrogenase (EC 1.2.1.2) - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 23-Sep-2002
C:Accession: A47117
R:Chow, C.M.; Rajbhandary, U.L.
J. Bacteriol. 175, 3703-3709, 1993
A:Title: Developmental regulation of the gene for formate dehydrogenase in Neurospora cr
A:Reference number: A47117; MUID:93285982; PMID:8509325
A:Accession: A47117
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-375 <CHO>
A:Cross-references: GB:L13964; NID:g293949; PIDN:AAA99900.1; PID:g1321604
C:Genetics:
A:Introns: 1/3; 17/3
C:Superfamily: Neurospora formate dehydrogenase
C:Keywords: oxidoreductase

Query Match 2.7%; Score 11; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 TAGISDHYDL 130
DB 91 TAGISDHYDL 101

RESULT 9

actA protein - *Emeticella nidulans*
C:Species: *Emeticella nidulans*, *Aspergillus nidulans*
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Sep-2002
C:Accession: S30088

R:Saleeba, J.A.; Cobbett, C.S.; Hynes, M.J.

Mol. Gen. Genet. 235, 349-358, 1992

A:Title: Characterization of the *actA* gene of *Aspergillus nidulans*.

A:Reference number: S30088; MUID:93101140; PMID:1465107

A:Accession: S30088

A:Molecule type: DNA

A:Residues: 1-377 <SAL>

A:Cross-references: EMBL:Z11612; NID:g5544; PIDN:CAA7687.1; PID:g5545

C:Genetics:

A:Gene: *actA*

A:Map position: 1

A:Introns: 5/3; 64/2

C:Superfamily: Neurospora formate dehydrogenase

Query Match 2.7%; Score 11; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10

formate dehydrogenase (FDH) - *Arabidopsis thaliana*
T51423
N:Alternate names: protein T9L3_80
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 23-Sep-2002
C:Accession: T51423
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.;
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25394
A:Accession: T51423
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-384 <SAT>
A:Cross-references: EMBL:AL391149
A:Experimental source: cultivar Columbia; BAC clone T9L3
C:Genetics:

A:Map position: 5

A:Introns: 29/3; 88/1; 268/2; 347/3

A>Note: T9L3 80

C:Superfamily: Neurospora formate dehydrogenase

Query Match 2.7%; Score 11; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 313 QPAPKDPHPR 323
DB 318 QPAPKDPHPR 328

RESULT 11

conserved hypothetical protein [imported] - *Sulfolobus solfataricus*
C:Species: *Sulfolobus solfataricus*
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: D90271
R:She, O.; Sindh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aveyez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.U.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A99139
A:Accession: D90271
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-89 <RUR>
A:Cross-references: GB:AE006641; NID:g13814371; PIDN:AAK41427.1; GSPDB:GN00155
C:Genetics:

A:Gene: SSO7348

Query Match 2.0%; Score 8; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 197 TVAAGRIG 204
DB 19 TVAAGRIG 26

RESULT 12

hypothetical protein Atu0077 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

C:Accession: AH2585
R:Wood, D.W.; Sebhal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, H.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutlyavin, T.; Levy, R.; Li, M.; Mclell
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.

A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; MUID:21608550; PMID:11743193

A:Accession: AH2585
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <KUR>
A:Cross-references: GB:AE008688; PIDN:AL41102.1; PID:g17738393; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:

A:Gene: Atu0077
A:Map position: circular chromosome
C:Superfamily: phosphoglycerate dehydrogenase

Query Match 2.0%; Score 8; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 YIVNTARG 286
DB 240 YIVNTARG 247

RESULT 13

probable dehydrogenase [imported] - Agrobacterium tumefaciens (strain C58, Cereon)

C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

C:Accession: G97367
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A:Title: Genome sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: G97367
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK85896.1; PID:g15154941; GSPDB:GN00169
C:Genetics:

A:Gene: AGR_C_114
A:Map position: circular chromosome
C:Superfamily: phosphoglycerate dehydrogenase

Query Match 2.0%; Score 8; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 279 YIVNTARG 286
DB 240 YIVNTARG 247

RESULT 14
E89779
NAD-dependent formate dehydrogenase [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: E89779
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cul, L.; O
ma, A.; Mizutani-Ul, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: E89779
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-374 <KUR>
A:Cross-references: GB:BA000018; PID:g13700093; PIDN:BA01392.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:

A:Gene: fdh

Query Match 2.0%; Score 8; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 319 DHPWRTMP 326
DB 311 DHPWRTMP 318

RESULT 15

hypothetical protein RV0106 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: B70752
R:Coile, S.T.; Davies, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
Connor, R.; Blythe, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98293987; PMID:9634230

A:Accession: B70752

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-398 <COL>

A:Cross-references: GB:Z74410; GB:AL123456; NID:g3261600; PIDN:CAA88919.1; PID:g14057

A:Experimental source: strain H37RV
C:Genetics:

A:Superfamily: conserved hypothetical protein yjc

Query Match 2.0%; Score 8; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 AVLRRLAP 213
DB 181 AVLRRLAP 188

Search completed: July 24, 2003, 21:02:09
Job time: 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:50:03 ; Search time 22 Seconds

(without alignments)
857.169 Million cell updates/sec

Title: US-09-996-008B-2
Perfect score: 401
Sequence: 1 MAKVLCVLYDDPDVGYPKTY.....YSKGNATGSGSEAKFKKAV 401

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size: 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	339	84.5	400	1 FDH_PSES	P33160 pseudomonas
2	16	4.0	361	1 FDH_PICAN	P33677 pichia angu
3	16	4.0	377	1 FDH_HORVU	O92r18 hordeum vul
4	12	3.0	381	1 FDH_SOLTU	O07511 solanum tub
5	11	2.7	375	1 FDH_NEUCR	O07103 neurospora
6	11	2.7	376	1 FDH_ORYSA	O98xp2 oryza sativ
7	11	2.7	377	1 FDH_EMENT	O08134 emeritella
8	11	2.7	384	1 FDH_ARATH	O95764 arabidopsis
9	8	2.0	190	1 FRDA_DROME	O95385 drosophila
10	8	2.0	398	1 Y106_MYCTU	O10899 mycobacteri
11	8	2.0	430	1 PURA_XANAC	O8pnb5 xanthomonas
12	8	2.0	430	1 PURA_XANAC	O8pnb5 xanthomonas
13	7	1.7	126	1 PRDB_SULSO	O94xb8 sulfobus
14	7	1.7	134	1 Y605_METUJ	O96022 methanococ
15	7	1.7	161	1 UCNB_HUMAN	O96963 homo sapien
16	7	1.7	176	1 DSBH_ECOLI	P59343 escherichia
17	7	1.7	176	1 DSBH_ECOLI	P30018 escherichia
18	7	1.7	176	1 DSBH_ECOLI	O34135 shigella fl
19	7	1.7	206	1 COAE_HAETN	P44920 haemophilus
20	7	1.7	211	1 RL1_HALCV	P05966 halobacteri
21	7	1.7	211	1 RL1_HALCV	P15757 halobacteri
22	7	1.7	277	1 MURI_BRUME	O8ybk2 bruceella me
23	7	1.7	287	1 DTC_HUMAN	O94bx3 homo sapien
24	7	1.7	293	1 RBSK_BACSU	P36945 bacillus su
25	7	1.7	309	1 DHT2_LACCO	P14295 lactobacilli
26	7	1.7	314	1 PIX2_BRARE	O9w522 brachydanio
27	7	1.7	314	1 RLAO_THEAC	P57692 thermoplas
28	7	1.7	329	1 YJ31_ARCFU	O28149 archaeoglob
29	7	1.7	334	1 YEBA_SCHPO	O14075 schizosacch
30	7	1.7	350	1 YN14_YEAST	P35833 saccharomyc
31	7	1.7	391	1 CML1_YEAST	P38101 saccharomyc
32	7	1.7	394	1 UDG_STRPN	O57346 streptococc
33	7	1.7	398	1 HEM1_METTM	P42809 methanobact

34	7	1.7	402	1 HEM1_METTH	O27093 methanobact
35	7	1.7	402	1 OPS4_CANAL	P46596 candida alb
36	7	1.7	403	1 KICS_MOUSE	P19901 mus musculu
37	7	1.7	408	1 PEPT_CLOAB	O97188 clostridium
38	7	1.7	414	1 SVS2_RAT	P22006 rattus norv
39	7	1.7	418	1 HLT_VIBPA	O99289 vibrio para
40	7	1.7	427	1 FLTL_HUMAN	O75955 homo sapien
41	7	1.7	428	1 FLTL_MOUSE	O08917 mus musculu
42	7	1.7	428	1 FLTL_RAT	O92161 rattus norv
43	7	1.7	463	1 GUN_BACSP	P28019 bacillus sp
44	7	1.7	464	1 KTR4_YEAST	P38131 saccharomyc
45	7	1.7	486	1 MENE_BACSU	P23971 bacillus su

ALIGNMENTS

RESULT 1	ID	FDH_PSES	STANDARD	PRT	400 AA.
AC	P33160	FDH_PSES	STANDARD	PRT	400 AA.
DT	01-OCT-1993 (Rel. 27, Created)				
DT	01-NOV-1995 (Rel. 32, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate dehydrogenase) (FDH).				
DE	Pseudomonas sp. (strain 101) (Achromobacter parvulus T1).				
OC	Bacteria; Proteobacteria.				
OX	NCBI_TaxID=33067;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=92063805; PubMed=1954846;				
RA	Tishkov V.I., Gaikina A.G., Egorov A.M.;				
RT	"NAD-dependent formate dehydrogenase of methylotrophic bacteria				
RT	Pseudomonas sp. 101: cloning, expression, and study of the genetic				
RT	structure.";				
RL	Dokl. Akad. Nauk SSSR 317:745-748(1991).				
RN	[2]				
RP	SEQUENCE OF 1-393.				
RX	MEDLINE=90290536; PubMed=2357236;				
RA	Popov V.O., Shumilina I.A., Uselinskaya T.B., Lamzin V.S., Egorov T.A.;				
RT	"NAD-dependent formate dehydrogenase from methylotrophic bacteria				
RT	Pseudomonas sp. 101. I. Amino acid sequence.";				
RL	Bioorg. Khim. 16:324-335(1990).				
RN	[3]				
RP	X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).				
RX	MEDLINE=92283270; PubMed=1597184;				
RA	Lamzin V.O., Alekhin A.E., Stokopytov B.V., Yukhnovich M.G.,				
RT	Popov V.O., Harutyunyan E.H., Wilson K.S.;				
RT	"Crystal structure of NAD-dependent formate dehydrogenase.";				
RL	Eur. J. Biochem. 206:441-452(1992).				
RN	[4]				
RP	MUTAGENESIS OF CYS-255.				
RX	MEDLINE=93249485; PubMed=8484798;				
RA	Tishkov V.I., Gaikina A.G., Marchenko G.N., Egorova O.A., Sheluhov D.V.,				
RT	Kulikova L.B., Dementieva L.A., Egorov A.M.;				
RT	"Catalytic properties and stability of a Pseudomonas sp.101 formate				
RT	dehydrogenase mutants containing Cys-255-Ser and Cys-255-Met				
RT	replacements.";				
RL	Biochem. Biophys. Res. Commun. 192:976-981(1993).				
CC	"- CATALYTIC ACTIVITY: Formate + NAD(+) - CO(2) + NADH.				
CC	"- SUBUNIT: Homodimer				
CC	"- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID				
CC	DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.				
DR	PIR: J00334; J00334.				
DR	PDB: 2NAD; 26-JAN-95.				
DR	InterPro: IPR006139; 2-Hacid_DH.				
DR	InterPro: IPR006140; 2-Hacid_DH.C.				
DR	Pfam: PF00389; 2-Hacid_DH.C.				
DR	Pfam: PF02826; 2-Hacid_DH.C.				
DR	PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.				
DR	PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.				

DR PROSITE; PS00671; D_2-HYDROXYACID_DH_3; 1.
 KW Oxidoreductase; NAD; 3D-structure.
 FT INIT_MET 0 0
 FT DOMAIN 1 146 CATALYTIC.
 FT DOMAIN 147 333 COENZYME-BINDING.
 FT DOMAIN 334 400 CATALYTIC.
 FT NP_BIND 192 226 NAD (ADP PART).
 FT ACT_SITE 5 5 POTENTIAL.
 FT ACT_SITE 284 284 SUBSTRATE BINDING.
 FT MUTAGEN 255 255 C->S,M: HIGH RESISTANCE TO INACTIVATION
 BY HG(2+); INCREASED STABILITY AT 25 C
 AND DECREASED THERMOSTABILITY AT 45 C.
 FT CONFLICT 77 77 D->S (IN REF. 1).
 FT CONFLICT 138 139 TV->VT (IN REF. 1).
 FT CONFLICT 145 145 C->V (IN REF. 1).
 FT CONFLICT 215 216 VH->HV (IN REF. 1).
 FT CONFLICT 327 327 N->D (IN REF. 1).
 FT STRAND 2 6
 FT STRAND 8 8
 FT TURN 12 13
 FT TURN 31 32
 FT TURN 46 47
 FT STRAND 52 52
 FT TURN 53 55
 FT TURN 56 58
 FT HELIX 60 65
 FT TURN 66 67
 FT STRAND 69 73
 FT TURN 79 80
 FT HELIX 82 87
 FT TURN 88 89
 FT STRAND 92 96
 FT TURN 97 98
 FT STRAND 99 99
 FT HELIX 103 103
 FT TURN 105 110
 FT TURN 112 113
 FT TURN 116 119
 FT STRAND 125 126
 FT STRAND 127 127
 FT HELIX 129 134
 FT TURN 135 136
 FT STRAND 138 141
 FT TURN 143 144
 FT HELIX 147 162
 FT TURN 163 164
 FT HELIX 165 173
 FT TURN 174 175
 FT HELIX 179 183
 FT TURN 184 185
 FT TURN 190 191
 FT STRAND 193 197
 FT HELIX 201 210
 FT HELIX 211 213
 FT TURN 214 214
 FT STRAND 216 220
 FT HELIX 227 233
 FT TURN 234 234
 FT STRAND 235 237
 FT HELIX 241 244
 FT STRAND 245 247
 FT TURN 250 253
 FT TURN 259 263
 FT STRAND 265 265
 FT HELIX 267 270
 FT TURN 271 272
 FT TURN 275 276
 FT STRAND 277 281
 FT HELIX 285 287
 FT STRAND 288 288
 FT HELIX 290 298
 FT TURN 299 300
 FT STRAND 302 307

FT TURN 317 318
 FT HELIX 320 322
 FT TURN 323 323
 FT TURN 325 326
 FT STRAND 327 327
 FT HELIX 334 336
 FT HELIX 338 357
 FT TURN 358 358
 FT HELIX 363 365
 FT STRAND 366 368
 FT STRAND 373 373
 FT TURN 375 375
 FT HELIX 376 380
 SQ SEQUENCE 400 AA; 44005 MW; A354151D4312991F CRC64;
 Query Match 84.5%; Score 339; DB 1; Length 400;
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 63 YLESNHGLVVTSDKDGPDSEFERELVDADVVISOPFWAYLTPERIARAKNLKALTAG 122
 DB 62 YLESNHGLVVTSDKDGPDSEFERELVDADVVISOPFWAYLTPERIARAKNLKALTAG 121
 QY 123 IGSDDVDLQSAIDRNVTVAEVTCNSISVAEHVMMILSLVRNYLPSHEARKGNIIAD 182
 DB 122 IGSDDVDLQSAIDRNVTVAEVTCNSISVAEHVMMILSLVRNYLPSHEARKGNIIAD 181
 QY 183 CVSHAYDLEAMHYGTVAAGRIGLAVLRLLAPFVHLHYDRHLPSVEKELNLTWHATR 242
 DB 182 CVSHAYDLEAMHYGTVAAGRIGLAVLRLLAPFVHLHYDRHLPSVEKELNLTWHATR 241
 QY 243 EDMYPCVDVVTLCPLHPETEHMINDETLKLFRGAYIVNTANGKLCDDRAVAIALESGR 302
 DB 242 EDMYPCVDVVTLCPLHPETEHMINDETLKLFRGAYIVNTANGKLCDDRAVAIALESGR 301
 QY 303 LAGYAGDVPWPAPADHPRTMPYNGMTHIGTILTAQARAAGTRLECEFFGGRPI 362
 DB 302 LAGYAGDVPWPAPADHPRTMPYNGMTHIGTILTAQARAAGTRLECEFFGGRPI 361
 QY 363 RDEYLIYOGGALAGTGAHSYSKGNATGSGSEAAKFKKAV 401
 DB 362 RDEYLIYOGGALAGTGAHSYSKGNATGSGSEAAKFKKAV 400
 RESULT 2
 FDH_PICAN STANDARD; PRT; 361 AA.
 AC P33677;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
 dehydrogenase) (FDH).
 GN PMDH.
 OS Pichia angusta (Yeast) (Hansenula polymorpha).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Pichia.
 OX NCBI_TaxID=4905;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hollenberg C.P., Janowicz Z.;
 RT "DNA-molecules coding for FMDH control regions and structured gene for
 a protease having FMDH-activity and their uses.";
 RL Patent number EP0299108, 18-JAN-1989.
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: A06214; CA00531.1; -.
 DR HSSP: P33160; ZNAD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH.C.
 DR Pfam: PF02826; 2-Hacid_DH.C.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 KW Oxidoreductase; NAD.
 FT INT_MET 0 PROBABLE.
 FT NP_BIND 164 199 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 257 257 SUBSTRATE BINDING (BY SIMILARITY).
 SQ SEQUENCE 361 AA; 39779 MW; 0FC001366F9E479B CRC64;
 Query Match 4.0%; Score 16; DB 1; Length 361;
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 308 GDVWFPPQAPKDPHMR 323
 Db 280 GDVWFPPQAPKDPHMR 295
 RESULT 3
 FDH_HORVU STANDARD; PRT; 377 AA.
 ID FDH_HORVU
 AC 09ZRI8;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
 DE dependent formate dehydrogenase) (FDH).
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticaceae; Hordeum.
 OX NCBI_TaxID=4513;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Ehimedaka No.1; TISSUE=Root;
 RX MEDLINE=98150277; PubMed=9489019;
 RA Suzuki K., Itai R., Suzuki K., Nakanishi H., Nishizawa N.K.,
 RA Yoshimura E., Mori S.;
 RT "Formate dehydrogenase, an enzyme of anaerobic metabolism, is induced
 RT by iron deficiency in barley roots.";
 RL Plant Physiol. 116:725-732(1998).
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D88272; BA36181.1; -.
 DR HSSP: P33160; ZNAD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH.C.
 DR Pfam: PF02826; 2-Hacid_DH.C.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; FALSE_NEG.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.

KW Oxidoreductase; NAD; Mitochondrion; Transit peptide.
 FT TRANSIT 1 21. MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 22 377. FORMATE DEHYDROGENASE.
 FT NP_BIND 191 225 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 283 283 SUBSTRATE BINDING (BY SIMILARITY).
 SQ SEQUENCE 377 AA; 41545 MW; E28C1FE24E9225C7 CRC64;
 Query Match 4.0%; Score 16; DB 1; Length 377;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 308 GDVWFPPQAPKDPHMR 323
 Db 306 GDVWFPPQAPKDPHMR 321
 RESULT 4
 FDH_SOLTU STANDARD; PRT; 381 AA.
 ID FDH_SOLTU
 AC 007511; 09ZR28;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
 DE dependent formate dehydrogenase) (FDH).
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. BF15;
 RA Hourton-Cabassa C., Ambard-Bretteville F., Remy R.,
 RA Colas des Francs-Small C.;
 RT "Evidence for multiple copies of formate dehydrogenase genes in
 RT plants: isolation of three potato fdh genes fdh1, fdh2 and fdh3.";
 RL (in) Plant Gene Register FGR98-102.
 RN
 RP SEQUENCE OF 3-381 FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=cv. BF15; TISSUE=Tuber;
 RX MEDLINE=94105343; PubMed=8278546;
 RA Colas des Francs-Small C., Ambard-Bretteville F., Small I.D.,
 RA Remy R.;
 RT "Identification of a major soluble protein in mitochondria from
 RT nonphotosynthetic tissues as NAD-dependent formate dehydrogenase.";
 RL Plant Physiol. 102:1171-1177(1993).
 RN
 RP REVISIONS TO N-TERMINUS.
 RC STRAIN=cv. BF15; TISSUE=Tuber;
 RA Colas des Francs-Small C.C.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE OF 26-54.
 RC STRAIN=cv. BF15; TISSUE=Tuber;
 RA Colas des Francs-Small C., Ambard-Bretteville F., Darpas A.,
 RA Sallantin M., Huet J.-C., Bernollet J.-C., Remy R.;
 RT "Variation of the polypeptide composition of mitochondria isolated
 RT from different potato tissues.";
 RL Plant Physiol. 98:273-278(1992).
 RN
 RP CHARACTERIZATION.
 RX PubMed=9490763;
 RA Hourton-Cabassa C., Ambard-Bretteville F., Moreau F.,
 RA Davy de Virville J., Remy R., Colas des Francs-Small C.;
 RT "Stress induction of mitochondrial formate dehydrogenase in potato
 RT leaves.";
 RL Plant Physiol. 116:627-635(1998).
 CC -1- FUNCTION: Involved in formate-dependent oxygen uptake coupled to
 CC ATP synthesis.
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer.

```

CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- TISSUE SPECIFICITY: Found at high levels in developing tubers, at
CC intermediate level in stems, veins, stolons, and stemens,
CC and at low level in leaves and roots.
CC -1- INDUCTION: Induced very rapidly by wounding, and slower by
CC darkness, chilling, drought, hypoxia, and treatments with formate,
CC abscisic acid, serine, sarcosine, pyruvate, acetate, ethanol or
CC methanol.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
CC -1- CAUTION: There are two other putative pseudogenes, fdh2 and fdh3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z99991; CAB17080.1; -
DR EMBL: Z99992; -; NOT_ANNOTATED_CDS.
DR EMBL: Z21493; CAA79702.2; -
DR PIR: J02272; J02272.
DR HSSP: P33160; 2NAD.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2-HYDROXYACID_DH.1; 1.
DR PROSITE: PS00670; D_2-HYDROXYACID_DH.3; 1.
DR PROSITE: PS00671; D_2-HYDROXYACID_DH.3; 1.
DR Oxidoreductase; NAD; Mitochondrion; Transit peptide.
KW TRANSIT 1 25 MITOCHONDRION.
FT CHAIN 1 26 381 FORMATE DEHYDROGENASE.
FT NP_BIND 195 229 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 287 287 SUBSTRATE BINDING (BY SIMILARITY).
SQ SEQUENCE 381 AA; 42038 MW; 920D351AC5A3A00E CRC64;

Query Match 3.0%; Score 12; DB 1; Length 381;
Best local Similarity 100.0%; Pred. No. 0.00037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 LTAGIGSDHVDL 130
DB 121 LTAGIGSDHVDL 132

```

RESULT 5

```

FDH_NEUCR STANDARD; PRT; 375 AA.
AC 007103;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-FEB-2003 (Rel. 42, Last annotation update)
DE Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
DE dehydrogenase) (FDH).
GN FDH OR 99H12.160.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=93385982; PubMed=8509325;
RA Chow C.M., Rajbandary U.L.,
RT "Developmental regulation of the gene for formate dehydrogenase in
RT Neurospora crassa."
RL J. Bacteriol. 175:3703-3709(1993).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;

```

```

RX PubMed=12655011;
RA Mannhaupt G., Montrone C., Haase D., Mewes H.-W., Aign V.,
RA Hohnselt J.D., Fartmann B., Nykatura G., Kempken F., Maier J.,
RA Schulte U.;
RT "What's in the genome of a filamentous fungus? Analysis of the
RT Neurospora genome sequence."
RL Nucleic Acids Res. 31:1944-1954(2003).
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- DEVELOPMENTAL STAGE: DEVELOPMENTALLY REGULATED. EXPRESSED ONLY
CC DURING CONIDINATION AND EARLY GERMINATION.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L13964; AAA99900.1; -
DR EMBL: AL451018; CAC18252.1; -
DR PIR: A47117; A47117.
DR HSSP: P33160; 2NAC.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR PROSITE: PS00065; D_2-HYDROXYACID_DH.1; 1.
DR PROSITE: PS00670; D_2-HYDROXYACID_DH.2; 1.
DR PROSITE: PS00671; D_2-HYDROXYACID_DH.3; 1.
DR Oxidoreductase; NAD.
KW NP_BIND 166 201 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 259 259 SUBSTRATE BINDING (BY SIMILARITY).
FT DOMAIN 364 374 ALA-RICH.
SQ SEQUENCE 375 AA; 40956 MW; 3073CB95FB204968 CRC64;

Query Match 2.7%; Score 11; DB 1; Length 375;
Best local Similarity 100.0%; Pred. No. 0.0038;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 TAGIGSDHVDL 130
DB 91 TAGIGSDHVDL 101

```

RESULT 6

```

FDH_ORYSA STANDARD; PRT; 376 AA.
AC Q95XP2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
DE dependent formate dehydrogenase) (FDH).
GN Oryza sativa (Rice).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN (1)
RP SEQUENCE FROM N.A.
RA Shiraiishi T., Fukusaki E., Kobayashi A.;
RT "NAD-dependent formate dehydrogenase."
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB019533; BAA77337.1; -
 DR HSSP: P33160; 2NAD.
 DR Gremene; Q9SKP2.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 DR Oxidoreductase; NAD; Mitochondrion; Transit peptide.
 FT TRANSIT 1 20 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 21 376 FORMATE DEHYDROGENASE.
 FT NP_BIND 190 224 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 282 282 SUBSTRATE BINDING (BY SIMILARITY).
 SQ SEQUENCE 376 AA; 41243 MW; 9D3964B64CF4040 CRC64;

Query Match 2.7%; Score 11; DB 1; Length 376;
 Best Local Similarity 100.0%; Pred. No. 0.0038;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPQAPK 318
 DB 305 GDVWFPQAPK 315
 |||||

RESULT 7
 FDH_EMENTI STANDARD: PRT; 377 AA.
 ID Q03134;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
 DE dehydrogenase) (FDH).
 GN ACIA.
 OS *Emergicella nidulans* (*Aspergillus nidulans*).
 CC Eukaryote; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 CC Eucotiales; Trichocomaceae; *Emergicella*.
 OX NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93101140; PubMed-1465107;
 RA Saleeba J.A., Cobbett C.S., Hynes M.J.;
 RT "Characterization of the amda-regulated acia gene of *Aspergillus*
 RT *nidulans*."; Mol. Gen. Genet. 235:349-358(1992).
 RL MoI. Gen. Genet. 235:349-358(1992).
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- INDUCTION: ACETATE INDUCTION MEDIATED BY AMDA REGULATORY GENE.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG. TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z11612; CAAT7687.1; -
 DR PIR: S30088; S30088.
 DR HSSP: P33160; 2NAD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.

DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; FALSE.NEG.
 KW Oxidoreductase; NAD.
 FT NP_BIND 161 191 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 271 271 SUBSTRATE BINDING (BY SIMILARITY).
 SQ SEQUENCE 377 AA; 41531 MW; 5219F5FEBCC931CE CRC64;

Query Match 2.7%; Score 11; DB 1; Length 377;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 120 TAGIGSDHVDL 130
 DB 86 TAGIGSDHVDL 96
 |||||

RESULT 8
 FDH_ARATH STANDARD: PRT; 384 AA.
 ID Q957E4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
 DE dependent formate dehydrogenase) (FDH).
 GN FDH1 OR FDH OR AT5G14780 OR T9L3_80.
 OS *Arabidopsis thaliana* (Mouse-ear cress).
 CC *Arabidopsis thaliana* (Mouse-ear cress).
 CC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; *Arabidopsis*.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nishikawa T., Fukusaki E., Kobayashi A.;
 RT "Formate dehydrogenase cDNA from *Arabidopsis thaliana*."; Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Li R., Bonham-Smith P.C., King J.;
 RT "Arabidopsis thaliana mRNA for NAD-dependent formate dehydrogenase
 RT 1."; Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 RL [3]
 RP SEQUENCE FROM N.A.
 RA Markwell J., Osterman J.C., Olson B.J., Skavdahl M., Ramberg H.,
 RA Germann M.C.;
 RT "Induction of leaf formate dehydrogenase by one-carbon metabolites."; Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
 RL [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE-21016721; PubMed-11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsuno M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Nanno K., Okumura S., Shilipo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stobbermann T., Pepin K., Spieth J., Sakhon M., Armstrong J., Becker M.,
 RA Belter E., Cordun H., Cordes M., Courtney B., Dente M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozeraky P., Riley A., Stromwall C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vill D., Baker J.,
 RA Kirchoff K., Roth K., King L., Bahret A., Miller B., Marra M.,
 RA Martensen R., McComble W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Wambolt R., Duesterhoelt A., Stiekema W., Pohl T.,
 RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grynopre B., Zimmermann W.,

RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirks W., Moolman P., Klein Lankhorst R.,
 RA Wetzinger T., Bothe G., Rose M., Haut J., Breier S.,
 RA Feldpausch M., Lamberth S., Villarejo R., Giesen J., Adles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.,
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 thaliana.";
 RL Nature 408:823-826(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, Columbia;
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT "Riken Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SGP consortium (Salk/Stanford/PGSC).";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CAPTARYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 DEHYDROGENASES FAMILY. STRONG, TO OTHER FGH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB023897; BAB8683.1; -
 DR EMBL: AF208028; AAF19435.1; -
 DR EMBL: AF208029; AAF19436.1; -
 DR EMBL: AF217195; AAF67100.1; -
 DR EMBL: AL391149; CAC01877.1; -
 DR EMBL: AY054285; AAL05944.1; -
 DR EMBL: AY039609; AAK62664.1; -
 DR EMBL: AY081734; AAL87387.1; -
 DR F1R; T51423; T51423.
 DR HSP; P33160; 2MAD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF02836; 2-Hacid_DH.C.1.
 DR PROSITE: PS00065; D-2-HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D-2-HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D-2-HYDROXYACID_DH_3; FALSE-NEG.
 KW Oxidoreductase; NAD; Mitochondrion; Transist peptide.
 FT TRANSIT 1 27 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 28 384 FORMATE DEHYDROGENASE.
 FT NP_BIND 198 232 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 290 290 SUBSTRATE BINDING (BY SIMILARITY).
 SQ SEQUENCE 384 AA; 42409 MW; A12BA423019D862B CRC64;
 Query Match 2.7%; Score 11; DB 1; Length 384;
 Best Local Similarity 100.0%; Pred. No. 0.0039;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 313 POPAPKDPWR 323
 Db 318 POPAPKDPWR 328
 RESULT 9
 FRDA_DROME STANDARD; PRT; 190 AA.
 ID FRDA_DROME
 AC Q9W385;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Frataxin homolog, mitochondrial precursor.
 FH GN CG8971.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20509989; PubMed=11054533;
 RA Canizares J., Blanca J.M., Navarro J.A., Monros E., Palau F.,
 RA Molto M.D.;
 RT "dth is a drosophila homologue of the Friedreich's ataxia disease
 RT gene";
 RL Gene 256:35-42(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang O., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Abghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolintsov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan Rocha S., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelle S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iqbal G.,
 RA Jaisai M., Kalush F., Kapran G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reineert K., Remington K., Saunders R.D.C., Scheetler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassenaar D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2165-2195(2000).
 CC -1- FUNCTION: PROBABLY INVOLVED IN IRON HOMEOSTASIS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE FRATAXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF208492; AAG35733.1; -
 DR EMBL: AF003446; AAF46449.1; -
 DR HSP; Q16595; 1DLX.
 DR F1R; F1R001794; Frataxin.
 DR InterPro: IPR002908; Frataxin_like.
 DR Pfam: PF01491; Frataxin_Cyay; 1.
 DR PRINTS; PR00904; FRATAXIN.

DR ProDom; PD238818; Frataxin_1like; 1.
 DR PROSITE; PS01344; FRATAXIN.1; 1.
 DR PROSITE; PS0810; FRATAXIN.2; 1.
 KW Mitochondrion; Transit peptide.
 FT TRANSIT 1 190 MITOCHONDRION.
 FT CHAIN ? 190 FRATAXIN HOMOLOG.
 SQ SEQUENCE 190 AA; 20921 MM; 8C48332C39F3036 CRC64;

Query Match 2.0%; Score 8; DB 1; Length 190;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 VGTVAAGR 202
 |||||
 DB 147 VGTVAAGR 154

RESULT 10
 ID Y106_MYCTU STANDARD; PRT; 398 AA.
 AC Q10899;

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein RV0106.
 GN RV0106 OR MT0115 OR MTCY251.25.
 OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;

RP SEQUENCE FROM N.A.

RA MEDLINE-96295987; PubMed-9634230; Garnier T., Churcher C., Harris D.,
 RA Cole S.T., Brosch R., Parkhill J., Garrier T., Barry C.E. III, Tekala F.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Ruster S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).

RP SEQUENCE FROM N.A.

RA STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weiman J., Knouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.

CC -1- SIMILARITY: SOME, TO P.DENTRIFICANS COB.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z74410; GAA98919.1; -
 DR EMBL; AE006923; AAK4337.1; -
 DR PIR; B70752; B70752.
 DR TIGR; MT0115; -

DR TubercuList; RV0106; -
 DR InterPro; IPR003495; COBW.
 DR Pfam; PF02492; COBW; 1.

KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 398 AA; 43699 MM; 2D75D9FBA167054A CRC64;

Query Match 2.0%; Score 8; DB 1; Length 398;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 AVLRLAP 213
 |||||
 DB 181 AVLRLAP 188

RESULT 11
 ID PURA_XANAC STANDARD; PRT; 430 AA.
 AC Q8PNS5;

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP-aspartate ligase)
 DE (AdSS) (AMPSase).
 GN PURA OR XAC1158.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=92829;

RP SEQUENCE FROM N.A.

RA STRAIN-306 / ATCC 13902 / XV 101;
 RA MEDLINE-22022145; PubMed-12024217;
 RA da Silva A.C.R., Ferro J.A., Relnach F.C., Farah C.S., Furian L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Camavari F., Cardozo J., Chamberg F., Clapina L.P.,
 RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Forthuglieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities."
 RL Nature 417:459-463(2002).

CC -1- FUNCTION: Plays an important role in the de novo pathway of purine
 CC nucleotide biosynthesis.
 CC -1- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
 CC adenylosuccinate.

CC -1- PATHWAY: AMP biosynthesis; first committed step.
 CC -1- SIMILARITY: Belongs to the adenylosuccinate synthetase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AE011745; AAM36030.1; -

DR HAMAP; MF_000111; -; 1.

DR InterPro; IPR001114; Asucc-synthetase.

DR Pfam; PF00709; Adenylosucc-synt; 1.

DR ProDom; PD001188; Asucc-synthetase; 1.

DR TIGRFAMS; TIGR00184; pura; 1.

DR PROSITE; PS01266; ADENYLOSUCCTIN_SYN_1; 1.

DR PROSITE; PS00513; ADENYLOSUCCTIN_SYN_2; 1.

DR Purine biosynthesis; Ligase; GTP-binding; Complete proteome.

FT NP_BIND 13 19

GTP (POTENTIAL).

FT ACT_SITE 141 141 BY SIMILARITY.
 FT ACT_SITE 148 148 BY SIMILARITY.
 SQ SEQUENCE 430 AA; 46199 MW; 770E439BEE1398F8 CRC64;

Query Match 2.08; Score 8; DB 1; Length 430;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 371 G GALACTG 378
 Db 247 G GALACTG 254

RESULT 12

PURA_XANCP
 ID PURA_XANCP STANDARD; PRT; 430 AA.
 AC 08P86;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)
 DE (AdSS) (Ampase).
 GN PURA OR XCC1053.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33913 / NCPPB 528;
 RX MEDLINE=2202145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Alves L.M.C., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Quesso L.M.C., do Amaral A.M., Bertolini M.C., Canaro L.E.A.,
 RA Ciccarelli G., Canavan F., Cardozo J., Chambergo F., Clapina L.P.,
 RA Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C.J., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Trufi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities."
 RL Nature 417:459-463(2002).
 CC -1- FUNCTION: Plays an important role in the de novo pathway of purine
 nucleotide biosynthesis.
 CC -1- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
 adenylosuccinate.
 CC -1- PATHWAY: AMP biosynthesis; first committed step.
 CC -1- SIMILARITY: Belongs to the adenylosuccinate synthetase family.
 CC -1- SIMILARITY: Belongs to the adenylosuccinate synthetase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AEO12202; AAM40352.1; -
 CC HAMAM; ME_00011; -; 1.
 DR InterPro: IPR001114; Asucc_synthase.
 DR Pfam; PF00709; Adenylosucc_synth; 1.
 DR PRODOM; PD001188; Asucc_synthase; 1.
 DR TIGRFAMs; TIGR00184; pura; 1.
 DR PROSITE; PS01266; ADENYLOSUCCLIN_SYN_1; 1.
 DR PROSITE; PS00513; ADENYLOSUCCLIN_SYN_2; 1.
 KM Purine biosynthesis; Ligase; GTP-binding; Complete proteome.

FT NP_BIND 13 19 GTP (POTENTIAL).
 FT ACT_SITE 141 141 BY SIMILARITY.
 FT ACT_SITE 148 148 BY SIMILARITY.
 SQ SEQUENCE 430 AA; 46163 MW; D36D57EAC2081B2D CRC64;

Query Match 2.08; Score 8; DB 1; Length 430;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 371 G GALACTG 378
 Db 247 G GALACTG 254

RESULT 13

PFDB_SUISO
 ID PFDB_SUISO STANDARD; PRT; 126 AA.
 AC Q90X88;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Prefoldin beta subunit (GimC beta subunit).
 GN PFDB OR SSO0730 OR C20_021.
 GN PFDB OR SSO0730 OR C20_021.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=20165948; PubMed=10701121;
 RA Charlebois R.L., Singh R.K., Chan-Welher C.C.-Y., Allard G., Chow C.,
 RA Confalonieri F., Curtis B., Duguet M., Eruso G., Paguy D.,
 RA Gasfardier T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
 RA Kushnava N., Lafleur E., Medina N., Peng X., Penny S.L., She O.,
 RA St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
 RA Regan M.A., Senses C.W.;
 RT "Gene content and organization of a 281-kbp contig from the genome of
 the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
 RL Genome 43:116-136(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=2133226; PubMed=11427726;
 RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Eruso G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Regan M.A., Senses C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -1- FUNCTION: Molecular chaperone capable of stabilizing a range of
 CC proteins. Seems to fulfill an ATP-independent, HSP70-like function
 CC in archaeal de novo protein folding (by similarity).
 CC -1- SUBUNIT: Heterohexamer of two alpha and four beta subunits (by
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE PREFOLDIN BETA SUBUNIT FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; Y18930; CAB57573.1; -
 DR EMBL; AEO06598; AAK41027.1; -
 DR PIR; D90221; D90221.
 DR HSP; O26774; LPK.

```

DR HAMAP: MF_00307; -: 1.
DR InterPro: IPR002777; PrefoldinK2.
DR Pfam: PF01920; KE2; 1.
KW Chaperone; Complete proteome.
SQ SEQUENCE 126 AA; 14543 MW; 5895DB741C8DE117 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 126;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 229 SVEKELN 235
DB 71 SVEKELN 77

RESULT 14
ID Y605_METJA STANDARD; PRT; 134 AA.
AC Q58022;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0605.
GN M0605.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;

RA SUTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOGYNE J.D.,
RA KETLAWE A.R., DOUGHERTY B.A., TOMB J.F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODOK A.,
RA SCOTT J.L., GEORGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KALNE B.P., BORODOVSKY M.,
RA KIENK H.-P., FRASER C.M., SMITH H.O., KOESE C.R., VENTER J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- SIMILARITY: STRONG. TO M.JANNASCHII M0142.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: U67508; AAB98599.1; -.
CC PIR: E64375; E64375.
CC TIGR: M0605; -.
CC InterPro: IPR002806; DUF103.
CC Pfam: PF01953; DUF103; 1.
CC ProDom: PD015886; DUF103; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 134 AA; 15596 MW; 4C1D5D528D3452C CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 134;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 291 RDAVARA 297
DB 37 RDAVARA 43

RESULT 15
UCN3_HUMAN

```

```

ID UCN3_HUMAN STANDARD; PRT; 161 AA.
AC 0969E3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urocortin III precursor (Ucn III) (Stresscopin).
DR UCN3 OR SPC.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;

RA LEWIS K., LI C., PERRIN M.H., BLOUNT A., KUNITAKE K., DONALDSON C.,
RA VAUGHAN J., REYES T.M., GULYAS J., FISCHER W., BILLETZKY J.L.,
RA RIVIER J., SAWCHENKO P.E., VALE W.W.;
RT "Identification of urocortin III, an additional member of the
RT corticotropin-releasing factor (CRF) family with high affinity for the
RT CRF2 receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7570-7575(2001).

RA Hsu S.Y., Hsueh A.J.W.;
RX MEDLINE=21227098; PubMed=11329063;
RT "Human stresscopin and stresscopin-related peptide are selective
RT ligands for the type 2 corticotropin-releasing hormone receptor."
RL Nat. Med. 7:605-611(2001).

CC -1- FUNCTION: Suppress food intake, delays gastric emptying and
CC decreases heat-induced edema. Might represent an endogenous ligand
CC for maintaining homeostasis after stress.
CC -1- SUBUNIT: Binds with high affinity to CRF receptors 2-alpha and 2-
CC beta.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE SAVVAGINE/CORTICOTROPIN-RELEASING
CC FACTOR/UCOTENSIN I FAMILY OF PEPTIDES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL: AF361943; AAK67317.1; -.
CC EMBL: AY026949; AAK11729.1; -.
CC MIM: 605901; -.
KW Hormone; Amidation; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 118
FT CHAIN 120 157 UROCORTIN III.
FT MOD_RES 157 157 AMIDATION (G-158 PROVIDE AMIDE GROUP).
SQ SEQUENCE 161 AA; 17862 MW; F92B2C8C3087857D CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 161;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 109 IAKAKNL 115
DB 137 IAKAKNL 143

Search completed: July 24, 2003, 20:59:36
Job time : 26 secs

```

RESULT 15
UCN3_HUMAN

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:56:19 ; Search time 99 Seconds
(without alignments)
1045.243 Million cell updates/sec

Title: US-09-996-008B-2
Perfect score: 401
Sequence: 1 MAKVLCVLYDDPVDGYPKTY.....YSKGNATGSGSEAKFKKAV 401

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	339	84.5	400	2	Q9R5V0
2	339	84.5	401	2	Q9R5V0 mycobacteri
3	31	7.7	402	2	Q9R5V1 mycobacteri
4	24	6.0	401	16	Q9R5V7 moraxella s
5	23	5.7	400	2	Q9R5V3 paracoccus
6	20	5.0	399	2	Q9R5V1 hyphomicrob
7	16	4.0	364	3	Q9R5V8 pinus pinas
8	16	4.0	364	3	Q9R5V8 candida met
9	16	4.0	364	3	Q9R5V8 candida boi
10	16	4.0	364	3	Q9R5V8 candida met
11	13	3.0	398	2	Q9R5V8 uncultured
12	13	3.0	398	2	Q9R5V8 uncultured
13	13	3.0	398	2	Q9R5V8 uncultured
14	11	2.7	376	3	Q9R5V1 saccharomyc
15	11	2.7	199	10	Q9R5V20 zeaxanthin
16	11	2.7	417	3	Q9R5V90 mycosphaere

17	8	2.0	89	17	Q9R5V3	Q9R5V3 sulfolobus
18	8	2.0	147	5	Q9R5V0	Q9R5V0 drosophila
19	8	2.0	164	2	Q9R5V1	Q9R5V1 thermus the
20	8	2.0	320	16	Q9R5V2	Q9R5V2 synechococc
21	8	2.0	334	16	Q9R5V6	Q9R5V6 agrobacteri
22	8	2.0	334	16	Q9R5V15	Q9R5V15 rhizobium m
23	8	2.0	345	12	Q9R5V3	Q9R5V3 anthracis a
24	8	2.0	374	16	Q9R5V4	Q9R5V4 streptococc
25	8	2.0	374	16	Q9R5V1	Q9R5V1 streptococc
26	8	2.0	382	2	Q9R5V5	Q9R5V5 pseudomonas
27	8	2.0	501	17	Q9R5V7	Q9R5V7 pyrobaculum
28	8	2.0	616	16	Q9R5V9	Q9R5V9 streptococc
29	8	2.0	617	16	Q9R5V1	Q9R5V1 streptococc
30	8	2.0	617	16	Q9R5V2	Q9R5V2 streptococc
31	8	2.0	617	16	Q9R5V8	Q9R5V8 streptococc
32	8	2.0	617	16	Q9R5V8	Q9R5V8 streptococc
33	8	2.0	618	16	Q9R5V4	Q9R5V4 streptococc
34	8	2.0	618	16	Q9R5V4	Q9R5V4 streptococc
35	8	2.0	618	16	Q9R5V7	Q9R5V7 streptococc
36	8	2.0	758	16	Q9R5V9	Q9R5V9 neisseria m
37	8	2.0	764	16	Q9R5V2	Q9R5V2 neisseria m
38	8	2.0	849	15	Q9R5V6	Q9R5V6 human immun
39	8	2.0	991	17	Q9R5V0	Q9R5V0 pyrobaculum
40	8	2.0	998	17	Q9R5V7	Q9R5V7 pyrobaculum
41	8	2.0	1105	16	Q9R5V3	Q9R5V3 caulobacter
42	8	2.0	1184	16	Q9R5V2	Q9R5V2 streptococc
43	7	1.7	40	11	Q9R5V6	Q9R5V6 ratius sp.
44	7	1.7	77	2	Q9R5V3	Q9R5V3 cornebacte
45	7	1.7	77	2	Q9R5V0	Q9R5V0 cornebacte

ALIGNMENTS

RESULT 1

ID	Q9R5V0	PRELIMINARY:	PRT:	400 AA.
AC	Q9R5V0:	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE	NAD(+)-dependent formate dehydrogenase (EC 1.2.1.2) (Fragment).			
OS	Mycobacterium vaccae.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OX	NCBI_TaxID=1810;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96158475; PubMed=8597552;			
RA	Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;			
RT	"Cloning of formate dehydrogenase gene from a methanol-utilizing			
RT	bacterium Mycobacterium vaccae N10.";			
RL	Appl. Microbiol. Biotechnol. 44:479-483(1995).			
DR	HSSP: P33160; 2NAD.			
DR	InterPro: IPR006139; 2-Hacid_DH.			
DR	InterPro: IPR006140; 2-Hacid_DH.C.			
DR	Pfam: PF00389; 2-Hacid_DH; 1.			
DR	Pfam: PF02826; 2-Hacid_DH.C; 1.			
DR	PROSITE: PS00065; D.2_HYDROXYACID_DH.1; 1.			
DR	PROSITE: PS00670; D.2_HYDROXYACID_DH.2; 1.			
DR	PROSITE: PS00671; D.2_HYDROXYACID_DH.3; 1.			
DR	SEQUENCE 400 AA; 44018 MW; 4422888CA22BD9B CRC64;			
QY	Query Match	84.5%; Score 339; DB 2; Length 400;		
QY	Best Local Similarity	100.0%; Pred. No. 0;		
QY	Matches	339; Conservative	0; Mismatches	0; Indels
QY			0; Gaps	0;
QY	63 YLSENGHLYVTSKDGDSVFERELVADAVYISQPEPAYITPERIAKAKLALTAG 122			
QY	62 YLSENGHLYVTSKDGDSVFERELVADAVYISQPEPAYITPERIAKAKLALTAG 121			
QY	123 IGSVDHDSADIRAVTAEVYCNISVAEIVVMILVANYLPSEHMAKGGMNAD 182			

```

Db      122  IGSDDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYPSEHARKGGMNIAD 181
Qy      183  CVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDVHLHYTRHRRLPESVEKELNTWHTATR 242
Db      182  CVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDVHLHYTRHRRLPESVEKELNTWHTATR 241
Qy      243  EDMYPVCDVYTLNCPRLPETEHMINDETLTKFRGAVIYVNTARGKLCDDRAVARALESGR 302
Db      242  EDMYPVCDVYTLNCPRLPETEHMINDETLTKFRGAVIYVNTARGKLCDDRAVARALESGR 301
Qy      303  LAGYAGDVWFPQAPKDPKHPRTMPYNGMTPHISGTTTLTAQARYAAGTREILECFEGRPI 362
Db      302  LAGYAGDVWFPQAPKDPKHPRTMPYNGMTPHISGTTTLTAQARYAAGTREILECFEGRPI 361
Qy      363  RDEYLIYOGGALAGTGAHSYSGKNATGSGSEAAKFKKAV 401
Db      362  RDEYLIYOGGALAGTGAHSYSGKNATGSGSEAAKFKKAV 400

```

RESULT 2

```

ID      093GV1  PRELIMINARY; PRT; 401 AA.
AC      093GV1;
DT      01-DEC-2001 (TReMBLrel. 19, Created)
DT      01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      Formate dehydrogenase (EC 1.2.1.2).
GN      FDH.
OS      Mycobacterium vaccae.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1810;
RN      [1]
RP      Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;
RT      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN      [2]
RX      SEQUENCE FROM N.A.
RA      MEDLINE=96158475; PubMed=8597552;
RA      Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;
RT      "Cloning of formate dehydrogenase gene from a methanol-utilizing
RT      bacterium Mycobacterium vaccae N10."
RL      Appl. Microbiol. Biotechnol. 44:479-483(1995).
DR      EMBL; AB072394; BAB69476.1; -.
DR      Interpro; IPR006139; 2-Hacid_DH.
DR      Interpro; IPR006140; 2-Hacid_DH_C.
DR      Pfam; PF00389; 2-Hacid_DH.1.
DR      Pfam; PF02826; 2-Hacid_DH_C.1.
DR      PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
DR      PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR      PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
KW      Oxidoreductase.
SQ      SEQUENCE 401 AA; 44149 MW; 8C4821B4790AD26 CRC64;

```

Query Match 84.5%; Score 339; DB 2; Length 401;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 339; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      63  YLESNHGLVVTSDKDDSPFERELVDADVISQPPWPAVLPERITAKAKNLKLTATAG 122
Db      63  YLESNHGLVVTSDKDDSPFERELVDADVISQPPWPAVLPERITAKAKNLKLTATAG 122
Qy      123  IGSDDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYPSEHARKGGMNIAD 182
Db      123  IGSDDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYPSEHARKGGMNIAD 182
Qy      183  CVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDVHLHYTRHRRLPESVEKELNTWHTATR 242
Db      183  CVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDVHLHYTRHRRLPESVEKELNTWHTATR 242
Qy      243  EDMYPVCDVYTLNCPRLPETEHMINDETLTKFRGAVIYVNTARGKLCDDRAVARALESGR 302
Db      243  EDMYPVCDVYTLNCPRLPETEHMINDETLTKFRGAVIYVNTARGKLCDDRAVARALESGR 302

```

```

Qy      303  LAGYAGDVWFPQAPKDPKHPRTMPYNGMTPHISGTTTLTAQARYAAGTREILECFEGRPI 362
Db      303  LAGYAGDVWFPQAPKDPKHPRTMPYNGMTPHISGTTTLTAQARYAAGTREILECFEGRPI 362
Qy      363  RDEYLIYOGGALAGTGAHSYSGKNATGSGSEAAKFKKAV 401
Db      363  RDEYLIYOGGALAGTGAHSYSGKNATGSGSEAAKFKKAV 401

```

RESULT 3

```

ID      008375  PRELIMINARY; PRT; 402 AA.
AC      008375;
DT      01-JUL-1997 (TReMBLrel. 04, Created)
DT      01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      NAD-dependent formate dehydrogenase (EC 1.2.1.2).
GN      FDH.
OS      Moraxella sp.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Moraxellaceae; Moraxella.
OX      NCBI_TaxID=479;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Galkin A.G.;
RT      Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR      EMBL; Y13245; CAAT73696.1; -.
DR      HSSP; P33160; 2NAD.
DR      Interpro; IPR006139; 2-Hacid_DH.
DR      Interpro; IPR006140; 2-Hacid_DH_C.
DR      Pfam; PF00389; 2-Hacid_DH.1.
DR      Pfam; PF02826; 2-Hacid_DH_C.1.
DR      PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
DR      PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR      PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
KW      NAD; Oxidoreductase.
SQ      SEQUENCE 402 AA; 44143 MW; C511CE1CE996BBB CRC64;

```

Query Match 7.7%; Score 31; DB 2; Length 402;
 Best Local Similarity 100.0%; Pred. No. 5; 5e-23;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      249  CDVYTLNCPRLPETEHMINDETLTKFRGAV 279
Db      249  CDVYTLNCPRLPETEHMINDETLTKFRGAV 279

```

RESULT 4

```

ID      0930E7  PRELIMINARY; PRT; 401 AA.
AC      0930E7;
DT      01-DEC-2001 (TReMBLrel. 19, Created)
DT      01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT      01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE      Probable NAD-dependent formate dehydrogenase.
GN      RA0251 OR SMA0478.
OS      Rhizobium meliloti (Sinorhizobium meliloti).
OC      Plasmid pSymA (megaplasmid 1).
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Rhizobiaceae; Sinorhizobium.
OX      NCBI_TaxID=382;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN=1021;
RA      MEDLINE=21396509; PubMed=11481432;
RA      Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA      Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA      Gurjal M., Hong A., Hutzar L., Hyman R.W., Kahn D., Kahn M.L.,
RA      Kaiman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA      Yeh K.-C., Davis R.W., Federici N.A., Long S.R.;
RT      "Nucleotide sequence and predicted functions of the entire

```

```

RT Shorhizobium meliloti pSyma megaplasmid."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL: AE007218; AAK64909.1; -.
DR InterPro: IPR006139; 2-Hacid_DH.C.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 401 AA; 43533 MW; C2693A417456204B CRC64;

Query Match
Best Local Similarity 100.0%; Score 24; DB 16; Length 401;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 GHTLVTSKDGPDSPFERELVDA 91
DB 70 GHTLVTSKDGPDSPFERELVDA 93

RESULT 5
OY 093GW3 PRELIMINARY; PRT; 400 AA.
AC 093GW3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NAD-dependent formate dehydrogenase.
GN Fdh.
OS Paracoccus sp. 12-A.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OX NCBI_TaxID=171192;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12-A;
RT Shioda T., Satoh T., Mineki S., Iida M., Taguchi H.;
RT "Cloning, nucleotide sequence and expression in Escherichia coli of
RT the gene for formate dehydrogenase of Paracoccus sp. 12-A, a formate-
RT assimilating bacterium.";
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB071373; BAB64941.1; -.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
SQ SEQUENCE 400 AA; 43757 MW; 13195AED9AC299D8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 23; DB 2; Length 400;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 378 GAHSYSGNATGSGSEAAKFKKA 400
DB 378 GAHSYSGNATGSGSEAAKFKKA 400

RESULT 6
OY 093UW1 PRELIMINARY; PRT; 399 AA.
AC 093UW1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE NAD+-dependent formate dehydrogenase.
GN Fdh.
OS Hyphomicrobium sp. JC17.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

```

```

OC Hyphomicrobiaceae; Hyphomicrobium.
OX NCBI_TaxID=142666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JC17;
RA Tanaka Y., Yoshida T., Watanabe K., Mitsunaga T.;
RT "Characterization, gene cloning and expression of NAD+-dependent
RT formate dehydrogenase from a methylotrophic bacterium Hyphomicrobium
RT sp. JC17.";
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB051073; BAB55449.1; -.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
SQ SEQUENCE 399 AA; 44044 MW; 853BAD2C635864DC CRC64;

Query Match
Best Local Similarity 100.0%; Score 20; DB 2; Length 399;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 249 CDVYTLNCPHPETHEMIND 268
DB 249 CDVYTLNCPHPETHEMIND 268

RESULT 7
OY 08VX85 PRELIMINARY; PRT; 248 AA.
AC 08VX85;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative NAD-dependent formate dehydrogenase (EC 1.2.1.2)
DE (Fragment).
OS Pinus pinaster (Maritime pine).
DE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root;
RA Dubos C., Plomion C.;
RT "Identification of water-deficit responsive genes in maritime pine
RT (Pinus pinaster Ait.) roots using non radioactive cDNA-AFLP.";
RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AJ300721; CAC83306.1; -.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
KW Oxidoreductase; NAD.
FT NON_TER 1
FT NON_TER 248
SQ SEQUENCE 248 AA; 27370 MW; 45E61CFADF55A0C8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 16; DB 10; Length 248;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 308 GDVWFPPQAPKDPHWR 323
DB 174 GDVWFPPQAPKDPHWR 189

RESULT 8
OY 013437 PRELIMINARY; PRT; 364 AA.
AC 013437;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)

```

```

DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE NAD-dependent formate dehydrogenase (EC 1.2.1.2).
GN FDH1 OR FDH3.
OS Candida boidinii (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
ON NCBI_TaxID=54777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=52;
RX MEDLINE=97369815; Pubmed=9226256;
RA Sakai Y., Murdanoto A.P., Konishi T., Iwamatsu A., Kato N.;
RT "Regulation of the formate dehydrogenase gene, FDH1, in the
RT methylotrophic yeast Candida boidinii and growth characteristics of an
RT FDH1-disrupted strain on methanol, methylamine, and choline.";
RL J. Bacteriol. 179:4480-4485(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC32195;
RA Siusarczyk H.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC32195;
RA Siusarczyk H.;
RL Theiss (1997), Heinrich-Heine-University Duesseldorf.
DR EMBL; AF004096; AAC49766.1; -.
DR EMBL; AJ245934; CAB54834.1; -.
DR HSSP; P33160; 2NAC.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH.C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
DR Oxidoreductase.
KW SEQUENCE 364 AA; 40370 MW; 1B30982E0D5B77E8 CRC64;

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 9
ID 000498 . PRELIMINARY; PRT; 364 AA.
AC 000498;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE NAD-dependent formate dehydrogenase.
OS Candida methylita.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
ON NCBI_TaxID=45353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96009885; Pubmed=7557425;
RA Allen S.J., Holbrook J.J.;
RT "Isolation, sequence and overexpression of the gene encoding NAD-
RT dependent formate dehydrogenase from the methylotrophic yeast Candida
RT methylita.";
RL Gene 162:99-104(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Hall L.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; X81129; CA57036.1; -.

```

```

DR HSSP; P33160; 2NAC.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH.C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
DR NAD.
KW SEQUENCE 364 AA; 40344 MW; D83DD1FA9D5EA303 CRC64;

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 10
ID 093968 PRELIMINARY; PRT; 364 AA.
AC 093968;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Formate dehydrogenase (EC 1.2.1.2).
GN FDH.
OS Candida boidinii (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
ON NCBI_TaxID=54777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCYC 1513;
RX MEDLINE=97369815; Pubmed=9226256;
RA Sakai Y., Murdanoto A.P., Konishi T., Iwamatsu A., Kato N.;
RT "Regulation of the formate dehydrogenase gene, FDH1, in the
RT methylotrophic yeast Candida boidinii and growth characteristics of an
RT FDH1-disrupted strain on methanol, methylamine, and choline.";
RL J. Bacteriol. 179:4480-4485(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NCYC 1513;
RA Labrou E.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ011046; CA09466.2; -.
DR HSSP; P33160; 2NAC.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH.C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00065; D_2_HYDROXYACID_DH_1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH_2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH_3; 1.
DR Oxidoreductase.
KW SEQUENCE 364 AA; 40266 MW; 34B727B7493D7B3F CRC64;

Query Match 4.0%; Score 16; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 308 GDVWFPQAPKDPHWR 323
DB 281 GDVWFPQAPKDPHWR 296

RESULT 11
ID 09F7P9 PRELIMINARY; PRT; 398 AA.
AC 09F7P9;
DT 01-MAR-2001 (TReMBLrel. 16, Created)

```

```
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
OS Predicted NAD-dependent formate dehydrogenase.
OC uncultured Proteobacterium EBAC31A08.
CX Bacteria; Proteobacteria; environmental samples.
RN NCBI_TaxID=133804;
RP [1]
RX MEDLINE=20446260; PubMed=10988064;
RA Beta O., Aravind L., Koonin E.V., Suzuki M.T., Hadd A., Nguyen L.P.,
RA Jovanovich S.B., Gates C.M., Feldman R.A., Spudich J.L., Spudich E.N.,
RA DeLong E.F.;
RT "Bacterial rhodopsin: evidence for a new type of phototrophy in the
RT sea.";
RL Science 289:1902-1906(2000).
DR EMBL AF279106; AAC104/0.1; -.
DR HSSP; P33160; 2NAD.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH_C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
SQ SEQUENCE 398 AA; 43992 MW; 7C2EC76FD5723C4C CRC64;

Query Match 3.2%; score 13; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 0.00022;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 GYADVWFPPQAP 317
Db 304 GYADVWFPPQAP 316
|||||
|||||

RESULT 12
Q08988 PRELIMINARY; PRT; 236 AA.
ID Q08988;
AC Q08988;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE ORF YPL275W.
GN YPL275W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Hebling U.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DE [2]
DE SEQUENCE FROM N.A.
RA Duesterhoeft A., Floeth M., Fritz M., Hilbert H., Moestl D.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA MIPs;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z73632; CAA98013.1; -.
DR HSSP; P33160; 2NAD.
DR InterPro; IPR006140; 2-Hacid_DH_C.
DR Pfam; PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
SQ SEQUENCE 236 AA; 26487 MW; 1C8EA0684606D166 CRC64;

Query Match 3.0%; score 12; DB 3; Length 236;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 QPARKDHPWRTM 325
|||||
|||||

DB 161 QPARKDHPWRTM 172

RESULT 13
Q08911 PRELIMINARY; PRT; 376 AA.
ID Q08911;
AC Q08911;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Chromosome XV reading frame ORF YOR388C.
GN Fdh1 OR YOR388C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H., Hebling U., Hofmann B.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPs;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z75296; CAA99720.1; -.
DR HSSP; P33160; 2NAD.
DR SGP; S0005915; Fdh1.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH_C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
SQ SEQUENCE 376 AA; 41714 MW; 67ECD86F9DDC2A02 CRC64;

Query Match 3.0%; score 12; DB 3; Length 376;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 QPARKDHPWRTM 325
Db 301 QPARKDHPWRTM 312
|||||
|||||

RESULT 14
Q08SC4 PRELIMINARY; PRT; 152 AA.
ID Q08SC4;
AC Q08SC4;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical conserved protein.
GN OB0717.
OS Oceanobacillus theyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
CX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus theyensis isolated from the Iheya
RT ridge and its unexpected adaptive capabilities to extreme
RT environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004595; BAC12673.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 152 AA; 17117 MW; 5C0BC4955C66D1D CRC64;

Query Match 2.7%; score 11; DB 16; Length 152;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 313 POPAPKDPWR 323
 |||||
 DB 82 POPAPKDPWR 92

RESULT 15

Q8W520 PRELIMINARY; PRT; 199 AA.
 ID Q8W520;
 AC Q8W520;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Formate dehydrogenase (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID:4577;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang H., Bohmert H.;
 RT "Plant one-carbon genes";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF439732; AAC33598.1;
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH_C.
 DR Pfam: PF02826; 2-Hacid_DH_1.
 DR Pfam: PF02826; 2-Hacid_DH_C; 1.
 DR PROSITE: PS00665; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 199 AA; 21445 MW; DFB9F8791861B6E9 CRC64;

Query Match 2.7%; Score 11; DB 10; Length 199;
 Best Local Similarity 100.0%; Pred. No. 0.013;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 308 GDVWFPQAPK 318
 |||||
 DB 188 GDVWFPQAPK 198

Search completed: July 24, 2003, 21:01:22
 Job time : 103 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:31:32 ; Search time 27 seconds
(without alignments)
628.395 Million cell updates/sec

Title: US-09-996-008B-2

Perfect score: 2135
Sequence: 1 MAKVLCVLYDDPVDGYPKTY.....YSKGNATGSEAKFKKAV 401

Scoring table:

BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUTS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	823	38.5	364	US-09-203-893A-2	Sequence 2, Appl1
2	820	38.4	364	US-09-203-893A-6	Sequence 6, Appl1
3	818	38.3	364	US-08-817-926-51	Sequence 51, Appl1
4	818	38.3	364	US-09-203-893A-32	Sequence 32, Appl1
5	817	38.3	364	US-09-203-893A-4	Sequence 4, Appl1
6	814	38.1	364	US-09-203-893A-16	Sequence 16, Appl1
7	813	38.1	364	US-09-203-893A-14	Sequence 14, Appl1
8	813	38.1	364	US-09-203-893A-18	Sequence 18, Appl1
9	811	38.0	364	US-09-203-893A-22	Sequence 22, Appl1
10	810	37.9	364	US-09-203-893A-20	Sequence 20, Appl1
11	810	37.9	364	US-09-203-893A-24	Sequence 24, Appl1
12	809	37.9	364	US-09-203-893A-12	Sequence 12, Appl1
13	808	37.8	364	US-09-203-893A-8	Sequence 8, Appl1
14	808	37.8	364	US-09-203-893A-10	Sequence 10, Appl1
15	808	37.8	364	US-09-203-893A-28	Sequence 28, Appl1
16	807	37.8	364	US-09-203-893A-26	Sequence 26, Appl1
17	807	37.8	364	US-09-203-893A-30	Sequence 30, Appl1
18	310	14.5	339	US-09-328-352-7424	Sequence 7424, Ap
19	309.5	14.5	345	US-09-222-817-2	Sequence 2, Appl1
20	309.5	14.5	345	US-09-222-786-2	Sequence 2, Appl1
21	309.5	14.5	530	US-09-222-817-12	Sequence 12, Appl1
22	309.5	14.5	530	US-09-222-786-12	Sequence 12, Appl1
23	309	14.5	539	US-09-134-001C-5118	Sequence 5118, Ap
24	307	14.4	530	US-09-222-817-14	Sequence 14, Appl1
25	307	14.4	530	US-09-222-786-14	Sequence 14, Appl1
26	283.5	13.3	375	US-09-252-991A-30745	Sequence 30745, A
27	276	12.9	328	US-08-956-012-1	Sequence 1, Appl1

28	265	12.4	439	US-08-553-999B-2	Sequence 2, Appl1
29	263.5	12.3	485	US-09-252-991A-25242	Sequence 25242, A
30	258	12.1	410	US-08-411-760-14	Sequence 14, Appl1
31	258	12.1	443	US-09-252-991A-27545	Sequence 27545, A
32	252	11.8	322	US-08-286-819A-2	Sequence 2, Appl1
33	252	11.8	322	US-08-980-357-2	Sequence 29, Appl1
34	252	11.8	2291	US-08-286-819A-29	Sequence 29, Appl1
35	252	11.8	2291	US-08-980-357-29	Sequence 29, Appl1
36	251.5	11.8	1113	US-09-252-991A-29215	Sequence 29215, A
37	249	11.7	439	US-09-328-352-6282	Sequence 6282, Ap
38	248.5	11.6	325	US-09-134-001C-5021	Sequence 5021, Ap
39	246.5	11.5	339	US-09-107-532A-7148	Sequence 7148, Ap
40	241	11.3	291	US-09-107-532A-5622	Sequence 5622, Ap
41	234.5	11.0	349	US-09-134-001C-4884	Sequence 4884, Ap
42	205	9.6	445	US-09-252-991A-20277	Sequence 20277, A
43	201	9.4	563	US-09-252-991A-24443	Sequence 24443, A
44	194	9.1	326	US-09-328-352-4608	Sequence 4608, Ap
45	191	8.9	356	US-09-107-532A-6936	Sequence 6936, Ap

ALIGNMENTS

RESULT 1
US-09-203-893A-2
; Sequence 2, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new gene sequences encoding these and use of
; TITLE OF INVENTION: the new Formate dehydrogenases
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203, 893A
; PRIOR FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
; US-09-203-893A-2

Query Match 38.5%, Score 823, DB 3, Length 364;

Best Local Similarity 45.1%, Pred. No. 3.5e-81;
Matches 171; Conservative 57; Mismatches 113; Indels 38; Gaps 6;

QY	3	KVLGVLYDDPVDGYPKTYARDLPKIDHYPGGQLIPPKAIDFTPGQLSGVSGELGRP	62	Sequence 2, Appl1
DB	2	KIVLVLYD-----AGKHADE-----EKLGGSTENKLGIAN	32	Sequence 1, Appl1
QY	63	YLESNGHTLVYSDKDPVSFERELVDADVVISQPEWPAFLPERIAKANLKLATAG	122	Sequence 2, Appl1
DB	33	WLKQGHLLITTSDEKGETSGLDKHHPADIIITTPHRAITTERLDKANKLKLAVAG	92	Sequence 1, Appl1
QY	123	IGSDHVDLQ--SAIDRNVYVAEVTYCNISVSAEHVVMILSLVRNYLPESHMARKGNI	180	Sequence 1, Appl1
DB	93	VGSIDHIDYDYNQGGKISVLEVGVSNVSAEHHVMTLVLRNFVAHQIINHDEW	152	Sequence 1, Appl1
QY	181	ADCSHVAYDEAMVGVPAAGRIGIAVLRRLAPRD-VILHTYDTRRLRDESEKELNLTWH	239	Sequence 1, Appl1
DB	153	AAIKADADIDEGKTATITAGRIGIRVLERLPNPKLLTYDQALKEAEVVGARRV	212	Sequence 1, Appl1
QY	240	ATREDMYPVCVDTVLNCPYLPHEPTEHMINDETLLFKRGAYLVNTARGCLCDRDVAAALE	299	Sequence 1, Appl1
DB	213	ENIEELVAQADIVVYVNPALHNGTKGLIKELLSFKGAWLVNTARGICVAAEDVAAALE	272	Sequence 1, Appl1
QY	300	SGRLAGTAGDVWPPQAPAKDHPWRTM-----PYNGMTPHISGTTLLAQAARYAAGTRETLE	354	Sequence 1, Appl1
DB	273	SGOLRGYCGDVWPPQAPAKDHPWDMRKYGAGNAMPYHSYGTTLDAQTRYAEGCTKNILE	332	Sequence 1, Appl1

QY 355 CFEGR-PIRDEYLIYOGG 372
 DB 333 SFTGKRDYRPDIILLNG 351

RESULT 2

US-09-203-893A-6
 : Sequence 6, Application US/09203893A
 : Patent No. 6242234
 : GENERAL INFORMATION:
 : APPLICANT: Kula, Maria-Regina
 : TITLE OF INVENTION: New mutants of Formate dehydrogenase from *Candida*
 : TITLE OF INVENTION: *boldini*, new gene sequences encoding these and use of
 : TITLE OF INVENTION: the new formate dehydrogenases
 : FILE REFERENCE: DEGU 3.0-012
 : CURRENT APPLICATION NUMBER: US/09/203,893A
 : PRIOR FILING DATE: 1998-12-02
 : PRIOR APPLICATION NUMBER: DE 19753350.7
 : PRIOR FILING DATE: 1997-12-03
 : NUMBER OF SEQ ID NOS: 37
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO: 6
 : LENGTH: 364
 : TYPE: PRT
 : ORGANISM: *Candida boldini*
 : US-09-203-893A-6

Query Match 38.4%; Score 820; DB 3; Length 364;
 Best Local Similarity 44.9%; Pred. No. 7.5e-81;
 Matches 170; Conservative 58; Mismatches 113; Indels 38; Gaps 6;

QY 3 KVLCLVLDPPVDGPKTYARDLPEKIDHYGGQILPTPKAIDFTPGQLGSVSGELGLRP 62
 DB 2 KIVLVLYD-----AGKHADE-----EKLYGATEKRLGIAN 32
 QY YLSENGHTLVVTSKDGSDSFERELVDADVISOFPWYALPPIERAKAKNLKLTAG 122
 DB 33 WLDQGHLLTTSDBESETSELDKHPDADIIITTPHPAYITKERLDKAKNLKLVVAG 92
 QY 123 IGSDDHDLQ--SAIDRNVTVAEYTYCNSISVAEHVMMILSYRNYLPSHEMARKGGMNI 180
 DB 93 VGSDDHDLQYINOTGKKSISLEVYGSNVVSAEYVMTLVLRNFPAPHEQIINHMEV 152
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVRLRLAPD-VHLHYTRHRLPESVEKELNLTWH 239
 DB 153 AATAKADYDIEGKTATIGAGRIGYRLELLEFPNRELLYYDYQALPKAEKEVGARV 212
 QY 240 ATREDMYPCDVVYTLNCPRLPETHMINDETLLKFRGAYIVNTARGKLCDDRAVARALE 299
 DB 213 ENIEELVAQADIVTVNAPRLHAGTKGLINKELLSFKKGAWLVNTARGAICVAEDVAAALE 272
 QY 300 SGRLAGYAGDVPFPPAPKDPHPRM-----PYNGMTPHISGTTTLTAQARYAAGTREILE 354
 DB 273 SGOIRGYGSDVWFPQAPKDPHPRMKNKYGAGNAMPHYSGTTLLDQTRYAAGTNYILE 332
 QY 355 CFEGR-PIRDEYLIYOGG 372
 DB 333 SFTGKRDYRPDIILLNG 351

RESULT 3

US-08-817-926-51
 : Sequence 51, Application US/08817926
 : Patent No. 6001590
 : GENERAL INFORMATION:
 : APPLICANT: Komeda, Toshihiro
 : APPLICANT: Suda, Hisako
 : APPLICANT: Tamai, Yukio
 : APPLICANT: Iwamatsu, Akihito
 : APPLICANT: Kato, No. 6001590uo
 : APPLICANT: Sakai, Yasuyoshi
 : TITLE OF INVENTION: PROMOTER/TERMINATOR FOR CANDIDA *BOLDINI*
 : TITLE OF INVENTION: FORMATE DEHYDROGENASE GENE

NUMBER OF SEQUENCES: 51
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/817,926
 FILING DATE: 09-MAY-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP96/02597
 FILING DATE: 12-SEP-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 234133/1995
 FILING DATE: 12-SEP-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 42536/1996
 FILING DATE: 29-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Bent, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 081356/0112
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)672-5300
 TELEFAX: (202)672-5399
 TELEX: 904136
 INFORMATION FOR SEQ ID NO: 51:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 364 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-817-926-51

Query Match 38.3%; Score 818; DB 3; Length 364;
 Best Local Similarity 44.9%; Pred. No. 1.2e-80;
 Matches 170; Conservative 57; Mismatches 114; Indels 38; Gaps 6;

QY 3 KVLCLVLDPPVDGPKTYARDLPEKIDHYGGQILPTPKAIDFTPGQLGSVSGELGLRP 62
 DB 2 KIVLVLYD-----AGKHADE-----EKLYGATEKRLGIAN 32
 QY YLSENGHTLVVTSKDGSDSFERELVDADVISOFPWYALPPIERAKAKNLKLTAG 122
 DB 33 WLDQGHLLTTSDBESETSELDKHPDADIIITTPHPAYITKERLDKAKNLKLVVAG 92
 QY 123 IGSDDHDLQ--SAIDRNVTVAEYTYCNSISVAEHVMMILSYRNYLPSHEMARKGGMNI 180
 DB 93 VGSDDHDLQYINOTGKKSISLEVYGSNVVSAEYVMTLVLRNFPAPHEQIINHMEV 152
 QY 181 ADCVSHAYDLEAMHVGTVAAAGRIGLAVRLRLAPD-VHLHYTRHRLPESVEKELNLTWH 239
 DB 153 AATAKADYDIEGKTATIGAGRIGYRLELLEFPNRELLYYDYQALPKAEKEVGARV 212
 QY 240 ATREDMYPCDVVYTLNCPRLPETHMINDETLLKFRGAYIVNTARGKLCDDRAVARALE 299
 DB 213 ENIEELVAQADIVTVNAPRLHAGTKGLINKELLSFKKGAWLVNTARGAICVAEDVAAALE 272
 QY 300 SGRLAGYAGDVPFPPAPKDPHPRM-----PYNGMTPHISGTTTLTAQARYAAGTREILE 354
 DB 273 SGOIRGYGSDVWFPQAPKDPHPRMKNKYGAGNAMPHYSGTTLLDQTRYAAGTNYILE 332
 QY 355 CFEGR-PIRDEYLIYOGG 372
 DB 333 SFTGKRDYRPDIILLNG 351

SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 22
 LENGTH: 364
 TYPE: PRT
 ORGANISM: Candida boldini
 US-09-203-893A-22

Query Match 38.0%; Score 811; DB 3; Length 364;
 Best Local Similarity 44.6%; Pred. No. 7.2e-80;
 Matches 169; Conservative 58; Mismatches 114; Indels 38; Gaps 6;

QY 3 KVLCLYDDPVGYPKTYARDLPKIDHYPGQILPFPKALDFTPGQLGSGVSGELGRP 62
 DB 2 KIVLVLD-----AGKHADE-----EKLYGATEKKGIAN 32
 QY 63 YLSENGHTLVVTSKDGSDSFERELVDADVVISOPFPAYLTPEBRIAKANKLALATAG 122
 DB 33 WLKDGHELTTSKEGETSELDKHIPDADIIITPPHAYITKERLDKANKLALVYAG 92
 QY 123 IGSDDHVDQ--SAIDRNVTAEVTCNSISVAEHVMMILSVNRYLPSEHMARKGNNI 180
 DB 93 VGSDDHIDLDYINQCKKISVLEVTGSNNVSAEHVMTMLVLRNFPVPAHQIINHMEV 152
 QY 181 ADCVSHAYDLAMHGVTAAGRIGLAVLRRLAPD-VHLHYTDHRLRPESVEKELNLTWH 239
 DB 153 AAIAKDAYDIEGKTATIGAGRIGRYLERLLPFPKELLYDYQALPKEAEKVGARRV 212
 QY 240 ATREDMPVCDVYVTLNCPHETEMINDELKLFKGAIVYNTARGLCDROAVARALE 299
 DB 213 ENIELVAQADIVTVNAPLAHGTGLINKELSKFKGAWLVNTARGAIVAEVAAALE 272
 QY 300 SGRLAGYAGDVWFPQAPAKDHPMRTM-----PYNGMTPHISGTTLTAQARYAAGRTRELE 354
 DB 273 SGOLRGYGDDWFPQAPAKDHPMRTMKNKYAGNAMPHYSGTTLLDQOTRYABEGTKNILE 332
 QY 355 CFPEGR-PIRDEYLVQGG 372
 DB 333 SFTGKFDYRPQDIILLNG 351

RESULT 10
 US-09-203-893A-20
 Sequence 20, Application US/09203893A
 Patent No. 6242234
 GENERAL INFORMATION:
 APPLICANT: Kula, Maria-Regina
 TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
 TITLE OF INVENTION: boldini, new Gene sequences encoding these and use of
 FILE REFERENCE: DECU 3.0-012
 CURRENT APPLICATION NUMBER: US/09/203, 893A
 PRIOR FILING DATE: 1998-12-02
 PRIOR APPLICATION NUMBER: DE 19753350.7
 NUMBER OF SEQ ID NOS: 37
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 20
 LENGTH: 364
 TYPE: PRT
 ORGANISM: Candida boldini
 US-09-203-893A-20

Query Match 37.9%; Score 810; DB 3; Length 364;
 Best Local Similarity 44.6%; Pred. No. 9.3e-80;
 Matches 169; Conservative 58; Mismatches 114; Indels 38; Gaps 6;

QY 3 KVLCLYDDPVGYPKTYARDLPKIDHYPGQILPFPKALDFTPGQLGSGVSGELGRP 62
 DB 2 KIVLVLD-----AGKHADE-----EKLYGATEKKGIAN 32
 QY 63 YLSENGHTLVVTSKDGSDSFERELVDADVVISOPFPAYLTPEBRIAKANKLALATAG 122
 DB 33 WLKDGHELTTSKEGETSELDKHIPDADIIITPPHAYITKERLDKANKLALVYAG 92

QY 123 IGSDDHVDQ--SAIDRNVTAEVTCNSISVAEHVMMILSVNRYLPSEHMARKGNNI 180
 DB 93 VGSDDHIDLDYINQCKKISVLEVTGSNNVSAEHVMTMLVLRNFPVPAHQIINHMEV 152
 QY 181 ADCVSHAYDLAMHGVTAAGRIGLAVLRRLAPD-VHLHYTDHRLRPESVEKELNLTWH 239
 DB 153 AAIAKDAYDIEGKTATIGAGRIGRYLERLLPFPKELLYDYQALPKEAEKVGARRV 212
 QY 240 ATREDMPVCDVYVTLNCPHETEMINDELKLFKGAIVYNTARGLCDROAVARALE 299
 DB 213 ENIELVAQADIVTVNAPLAHGTGLINKELSKFKGAWLVNTARGAIVAEVAAALE 272
 QY 300 SGRLAGYAGDVWFPQAPAKDHPMRTM-----PYNGMTPHISGTTLTAQARYAAGRTRELE 354
 DB 273 SGOLRGYGDDWFPQAPAKDHPMRTMKNKYAGNAMPHYSGTTLLDQOTRYABEGTKNILE 332
 QY 355 CFPEGR-PIRDEYLVQGG 372
 DB 333 SFTGKFDYRPQDIILLNG 351

RESULT 11
 US-09-203-893A-24
 Sequence 24, Application US/09203893A
 Patent No. 6242234
 GENERAL INFORMATION:
 APPLICANT: Kula, Maria-Regina
 TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
 TITLE OF INVENTION: boldini, new Gene sequences encoding these and use of
 FILE REFERENCE: DECU 3.0-012
 CURRENT APPLICATION NUMBER: US/09/203, 893A
 PRIOR FILING DATE: 1998-12-02
 PRIOR APPLICATION NUMBER: DE 19753350.7
 NUMBER OF SEQ ID NOS: 37
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 24
 LENGTH: 364
 TYPE: PRT
 ORGANISM: Candida boldini
 US-09-203-893A-24

Query Match 37.9%; Score 810; DB 3; Length 364;
 Best Local Similarity 44.6%; Pred. No. 9.3e-80;
 Matches 169; Conservative 58; Mismatches 114; Indels 38; Gaps 6;

QY 3 KVLCLYDDPVGYPKTYARDLPKIDHYPGQILPFPKALDFTPGQLGSGVSGELGRP 62
 DB 2 KIVLVLD-----AGKHADE-----EKLYGATEKKGIAN 32
 QY 63 YLSENGHTLVVTSKDGSDSFERELVDADVVISOPFPAYLTPEBRIAKANKLALATAG 122
 DB 33 WLKDGHELTTSKEGETSELDKHIPDADIIITPPHAYITKERLDKANKLALVYAG 92
 QY 123 IGSDDHVDQ--SAIDRNVTAEVTCNSISVAEHVMMILSVNRYLPSEHMARKGNNI 180
 DB 93 VGSDDHIDLDYINQCKKISVLEVTGSNNVSAEHVMTMLVLRNFPVPAHQIINHMEV 152
 QY 181 ADCVSHAYDLAMHGVTAAGRIGLAVLRRLAPD-VHLHYTDHRLRPESVEKELNLTWH 239
 DB 153 AAIAKDAYDIEGKTATIGAGRIGRYLERLLPFPKELLYDYQALPKEAEKVGARRV 212
 QY 240 ATREDMPVCDVYVTLNCPHETEMINDELKLFKGAIVYNTARGLCDROAVARALE 299
 DB 213 ENIELVAQADIVTVNAPLAHGTGLINKELSKFKGAWLVNTARGAIVAEVAAALE 272
 QY 300 SGRLAGYAGDVWFPQAPAKDHPMRTM-----PYNGMTPHISGTTLTAQARYAAGRTRELE 354
 DB 273 SGOLRGYGDDWFPQAPAKDHPMRTMKNKYAGNAMPHYSGTTLLDQOTRYABEGTKNILE 332
 QY 355 CFPEGR-PIRDEYLVQGG 372

Db 333 SFTGKFDYRPQDITLLNG 351

RESULT 12

US-09-203-893A-12

Sequence 12, Application US/09203893A
 Patent No. 6242234
 GENERAL INFORMATION:
 APPLICANT: Kula, Maria-Regina
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 FILE REFERENCE: DEGU 3.0-012
 CURRENT APPLICATION NUMBER: US/09/203,893A
 CURRENT FILING DATE: 1998-12-02
 PRIOR APPLICATION NUMBER: DE 19753350.7
 PRIOR FILING DATE: 1997-12-03
 NUMBER OF SEQ ID NOS: 37
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 12
 LENGTH: 364
 TYPE: PRM
 ORGANISM: Candida boidinii
 US-09-203-893A-12

Query Match 37.8%; Score 809; DB 3; Length 364;

Best Local Similarity 44.6%; Pred. No. 1.2e-79;
 Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

QY 3 KVLCLVLDPPVDCIPKTYARDLPRKIDHFGGOLPTPKAIDPTPGQLGSVSGELGRP 62
 DB 2 KIVLVLYD-----AGKHADE-----EKLYGCTENKIGIAN 32
 QY 63 YLESNGHTLVNTSDKOPDSVFERELVDADVVISOPWPAYLPERLAKKNTKLALTAG 122
 DB 33 WLDQGHLELTITSDKBESELDKHIPDADITTPHPAYITKERLDKKNKLKLVVAG 92
 QY 123 IGSDDHVDLQ--SAIDRNVTVAEVTYCNISVAEHVVMILSLVRYNLPSEHMAKGGWNI 180
 DB 93 VGSDDHIDLIDYINOTGKKISVLEVYGSNVSVAEHVVMTLVLRNFPAHEQIINHMEV 152
 QY 181 ADCVSHAYDLEAMHVGTVAAAGTGLAVRLRLAPD-VHLHYTRHRLPSEVEKELNLTWH 239
 DB 153 AAIAKDAYDIEGKTATIGAGRGYRLERLPEFNPELTYDYQALPKAEKVGARRV 212
 QY 240 ATREDMPYCDVYTNLCPLRPETEHMINDETLKLFRGAVIVMTARGKLCDDRAVARALE 299
 DB 213 ENIEELVAQDIYTVNAPRLHAGTKGLINKELLSKFKGAWLVNTARGAIVADVAALAE 272
 QY 300 SGRLAGYAGDVWFPQAPKDPHRTM---PYNGMTPHISGTTLPAAQARYAAGTREILE 354
 DB 273 SGOLRGYGGDVWFPQAPKDPHRTM---PYNGMTPHISGTTLPAAQARYAAGTREILE 332
 QY 355 CFPEGR-PIRDEYLYOQG 372
 DB 333 SFTGKFDYRPQDITLLNG 351

RESULT 13

US-09-203-893A-8

Sequence 8, Application US/09203893A
 Patent No. 6242234
 GENERAL INFORMATION:
 APPLICANT: Kula, Maria-Regina
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 FILE REFERENCE: DEGU 3.0-012
 CURRENT APPLICATION NUMBER: US/09/203,893A
 CURRENT FILING DATE: 1998-12-02
 PRIOR APPLICATION NUMBER: DE 19753350.7
 PRIOR FILING DATE: 1997-12-03

NUMBER OF SEQ ID NOS: 37
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 8
 LENGTH: 364
 TYPE: PRM
 ORGANISM: Candida boidinii
 US-09-203-893A-8

Query Match 37.8%; Score 808; DB 3; Length 364;

Best Local Similarity 44.6%; Pred. No. 1.5e-79;
 Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

QY 3 KVLCLVLDPPVDCIPKTYARDLPRKIDHFGGOLPTPKAIDPTPGQLGSVSGELGRP 62
 DB 2 KIVLVLYD-----AGKHADE-----EKLYGCTENKIGIAN 32
 QY 63 YLESNGHTLVNTSDKOPDSVFERELVDADVVISOPWPAYLPERLAKKNTKLALTAG 122
 DB 33 WLDQGHLELTITSDKBESELDKHIPDADITTPHPAYITKERLDKKNKLKLVVAG 92
 QY 123 IGSDDHVDLQ--SAIDRNVTVAEVTYCNISVAEHVVMILSLVRYNLPSEHMAKGGWNI 180
 DB 93 VGSDDHIDLIDYINOTGKKISVLEVYGSNVSVAEHVVMTLVLRNFPAHEQIINHMEV 152
 QY 181 ADCVSHAYDLEAMHVGTVAAAGTGLAVRLRLAPD-VHLHYTRHRLPSEVEKELNLTWH 239
 DB 153 AAIAKDAYDIEGKTATIGAGRGYRLERLPEFNPELTYDYQALPKAEKVGARRV 212
 QY 240 ATREDMPYCDVYTNLCPLRPETEHMINDETLKLFRGAVIVMTARGKLCDDRAVARALE 299
 DB 213 ENIEELVAQDIYTVNAPRLHAGTKGLINKELLSKFKGAWLVNTARGAIVADVAALAE 272
 QY 300 SGRLAGYAGDVWFPQAPKDPHRTM---PYNGMTPHISGTTLPAAQARYAAGTREILE 354
 DB 273 SGOLRGYGGDVWFPQAPKDPHRTM---PYNGMTPHISGTTLPAAQARYAAGTREILE 332
 QY 355 CFPEGR-PIRDEYLYOQG 372
 DB 333 SFTGKFDYRPQDITLLNG 351

RESULT 14

US-09-203-893A-10

Sequence 10, Application US/09203893A
 Patent No. 6242234
 GENERAL INFORMATION:
 APPLICANT: Kula, Maria-Regina
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 TITLE OF INVENTION: New mutants of formate dehydrogenase from Candida
 FILE REFERENCE: DEGU 3.0-012
 CURRENT APPLICATION NUMBER: US/09/203,893A
 CURRENT FILING DATE: 1998-12-02
 PRIOR APPLICATION NUMBER: DE 19753350.7
 PRIOR FILING DATE: 1997-12-03
 NUMBER OF SEQ ID NOS: 37
 SOFTWARE: Patentln Ver. 2.1
 SEQ ID NO 10
 LENGTH: 364
 TYPE: PRM
 ORGANISM: Candida boidinii
 US-09-203-893A-10

Query Match 37.8%; Score 808; DB 3; Length 364;

Best Local Similarity 44.6%; Pred. No. 1.5e-79;
 Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

QY 3 KVLCLVLDPPVDCIPKTYARDLPRKIDHFGGOLPTPKAIDPTPGQLGSVSGELGRP 62
 DB 2 KIVLVLYD-----AGKHADE-----EKLYGCTENKIGIAN 32
 QY 63 YLESNGHTLVNTSDKOPDSVFERELVDADVVISOPWPAYLPERLAKKNTKLALTAG 122

Db 33 WLKDGHELLITTSDEKSELTSDKHIPDADIIITTFPHPAYITTKERLDKANKLKVYAG 92
Qy 123 IGSDBHVDLQ--SADIRNTVAEVTYCNISVAEHVMMILSLVRYLPSHEMARGGNNI 180
Db 93 VGSDBHIDIDYINQGGKISVLEVTGSNVVAEHHVMTMLVNRNFPVPAHEDIIINHDEV 152
Qy 181 ADCSHAYDLBAMHVTVAAGRIGLAVLRRLAPFD-VHLHYTDHRRLPESVEKELNLTWH 239
Db 153 AAIKADAYDIEGKTATIGAGRIGRVLERLLPFKELLYDYDQALPKEAEKVGARRV 212
Qy 240 ATREDMPVCVVTINCLPHETEMINDETLKLFKCAIYVNTARGKLCDDRAVARALE 299
Db 213 ENIELVAQADIVTVNAPLHAGTKGLINKELLSKFKGAMLVNTARGAIVAAEDVAAALE 272
Qy 300 SGRLAGYAGDVWFPQAPKDPWRM-----PYNGMTPHISGTTTLTAQARYAGTRETLE 354
Db 273 SGRLRGYGGDVWFPQAPKDPWRMRRKYGAGNAMPHPYSGTTLDAOTRYAEGTKNILE 332
Qy 355 CFEEGR-PIRDEYLIVOGG 372
Db 333 SFTGKFDYRPQDITLLNG 351

RESULT 15

US-09-203-893A-28
; Sequence 28, Application US/09203893A
; Patent No. 6242234
; GENERAL INFORMATION:
; APPLICANT: Kula, Maria-Regina
; TITLE OF INVENTION: New mutants of Formate dehydrogenase from Candida
; TITLE OF INVENTION: boldin1, new Gene sequences encoding these and use of
; FILE REFERENCE: DEGU 3.0-012
; CURRENT APPLICATION NUMBER: US/09/203,893A
; CURRENT FILING DATE: 1998-12-02
; PRIOR APPLICATION NUMBER: DE 19753350.7
; PRIOR FILING DATE: 1997-12-03
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Candida boldin1
US-09-203-893A-28

Query Match 37.8%; Score 808; DB 3; Length 364;
Best Local Similarity 44.6%; Freq. No. 1.5e-79;
Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

Qy 3 KVLGVYDDPVDGYPKTYARDDLPKIDHYPGQILPTPKAIDFTPGOLLSGSGELGRP 62
Db 2 KIVLVLYD-----AGKHADE-----EKLXGVTENKLGIAN 32
Qy 63 YLENGHTLVVTSKDPDSVEFERELVDADVVISQFMPAYLTPERIAKANKLKLATAG 122
Db 33 WLKDGHELLITTSDEKSELTSDKHIPDADIIITTFPHPAYITTKERLDKANKLKVYAG 92
Qy 123 IGSDBHVDLQ--SADIRNTVAEVTYCNISVAEHVMMILSLVRYLPSHEMARGGNNI 180
Db 93 VGSDBHIDIDYINQGGKISVLEVTGSNVVAEHHVMTMLVNRNFPVPAHEDIIINHDEV 152
Qy 181 ADCSHAYDLBAMHVTVAAGRIGLAVLRRLAPFD-VHLHYTDHRRLPESVEKELNLTWH 239
Db 153 AAIKADAYDIEGKTATIGAGRIGRVLERLLPFKELLYDYDQALPKEAEKVGARRV 212
Qy 240 ATREDMPVCVVTINCLPHETEMINDETLKLFKCAIYVNTARGKLCDDRAVARALE 299
Db 213 ENIELVAQADIVTVNAPLHAGTKGLINKELLSKFKGAMLVNTARGAIVAAEDVAAALE 272
Qy 300 SGRLAGYAGDVWFPQAPKDPWRM-----PYNGMTPHISGTTTLTAQARYAGTRETLE 354
Db 273 SGRLRGYGGDVWFPQAPKDPWRMRRKYGAGNAMPHPYSGTTLDAOTRYAEGTKNILE 332

Qy 355 CFEEGR-PIRDEYLIVOGG 372
Db 333 SFTGKFDYRPQDITLLNG 351

Search completed: July 24, 2003, 20:40:28
Job time : 30 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:39:02 ; Search time 29 Seconds
(without alignments)
1642.161 Million cell updates/sec

Title: US-09-996-008B-2
Perfect score: 2135
Sequence: 1 MAKVLCVLYDDPDVGYPKTY.....YSKGNATGSGSEAKFKKAV 401

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues
Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2-6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2-6/ptodata/2/pubpaa/PC1_NEW_PUB.pep:*
3: /cgn2-6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2-6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2-6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2-6/ptodata/2/pubpaa/PC1US_PUBCOMB.pep:*
7: /cgn2-6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2-6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2-6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2-6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2-6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2-6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2-6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2-6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2-6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2-6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2-6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2-6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1428.5	66.9	387	15	US-10-156-761-9479 Sequence 9479, Ap
2	838.5	39.3	376	15	US-10-224-567-1 Sequence 1, Appl
3	781	36.6	343	9	US-09-815-242-12981 Sequence 12981, A
4	781	36.6	343	9	US-09-815-242-13150 Sequence 13150, A
5	780.5	36.6	335	9	US-09-815-242-5830 Sequence 5830, Ap
6	340.5	15.9	325	15	US-10-156-761-15009 Sequence 15009, A
7	310.5	14.5	530	10	US-09-738-626-4915 Sequence 4915, Ap
8	309.5	14.5	345	15	US-10-195-548-12 Sequence 12, Appl
9	309.5	14.5	530	15	US-10-195-548-12 Sequence 12, Appl
10	307	14.4	529	15	US-10-156-761-10267 Sequence 10267, A
11	297.5	13.9	410	9	US-09-815-242-11028 Sequence 11028, A
12	297	13.0	331	9	US-09-815-242-11761 Sequence 11761, A
13	277	12.9	445	15	US-10-177-293-92 Sequence 92, Appl
14	275.5	12.9	985	15	US-10-205-823-76 Sequence 76, Appl
15	264	12.4			

16	264	12.4	985	15	US-10-177-293-94	Sequence 94, Appl
17	258	12.1	510	9	US-09-815-242-10291	Sequence 10291, A
18	256	12.0	424	9	US-09-815-242-11557	Sequence 11557, A
19	254	11.9	427	9	US-09-815-242-13743	Sequence 13743, A
20	244.5	11.5	329	9	US-09-815-242-11811	Sequence 11811, A
21	234.5	11.0	335	15	US-10-156-761-9627	Sequence 9627, Ap
22	224.5	10.5	304	10	US-09-738-626-6082	Sequence 6082, Ap
23	217.5	10.2	333	9	US-09-815-242-10900	Sequence 10900, A
24	210.5	9.9	336	10	US-09-925-300-1408	Sequence 1408, Ap
25	210	9.8	325	15	US-10-122-466A-17	Sequence 17, Appl
26	206.5	9.7	315	10	US-09-922-501-14	Sequence 14, Appl
27	203.5	9.5	313	15	US-10-156-761-10272	Sequence 10272, A
28	195.5	9.2	319	15	US-10-156-761-12767	Sequence 12767, A
29	192	9.0	304	10	US-09-738-626-3580	Sequence 3580, Ap
30	126	5.9	171	9	US-09-939-980-307	Sequence 307, App
31	101.5	4.8	432	15	US-10-156-761-14180	Sequence 14180, A
32	95	4.4	444	15	US-10-156-761-13239	Sequence 13239, A
33	94	4.4	1144	15	US-10-156-761-13239	Sequence 7801, Ap
34	92	4.3	398	15	US-10-156-761-13711	Sequence 13711, A
35	91	4.3	327	11	US-09-284-320-16	Sequence 16, Appl
36	91	4.3	915	15	US-10-156-761-10050	Sequence 10050, A
37	90.5	4.2	566	9	US-09-815-242-13371	Sequence 13371, A
38	89.5	4.2	71	9	US-09-864-761-46245	Sequence 46245, A
39	87	4.1	792	9	US-09-815-242-11966	Sequence 11966, A
40	87	4.1	792	15	US-10-127-032-165	Sequence 165, App
41	87	4.1	1139	15	US-10-156-761-10856	Sequence 10856, A
42	87	4.1	1426	9	US-09-912-020-340	Sequence 340, App
43	86.5	4.1	3970	15	US-10-156-761-10429	Sequence 10429, A
44	86	4.0	716	15	US-10-156-761-10760	Sequence 10760, A
45	85.5	4.0	415	9	US-09-815-242-11084	Sequence 11084, A

ALIGNMENTS

RESULT 1
US-10-156-761-9479
Sequence 9479, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROKI
APPLICANT: SHIDA, TADAYOSHI
APPLICANT: SAKAKI, YOSHITUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9479
LENGTH: 387
TYPE: PRT
ORGANISM: Streptomyces avermiltis
US-10-156-761-9479
Query Match
Best Local Similarity 66.9%; Score 1428.5; DB 15; Length 387;
Matches 267; Conservative 42; Mismatches 74; Indels 1; Gaps 1;

QY 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGGQILTPKAIIDFTPOLGSGVSGEL 60
DB 1 MAKVLCVLYDDPDVGYPKTYARDDLPKIDHYPGGQILTPKAIIDFTPOLGSGVSGEL 60
QY 61 RPYLESNGHTLVVTSDDKDPDVSFERELVDADVVISQPFWPAVLTPERIAKNIKLAIT 120
DB 61 RPYLESNGHTLVVTSDDKDPDVSFERELVDADVVISQPFWPAVLTPERIAKNIKLAIT 120

QY 121 AGIGSDHVDLSAIDRNVTAEVTCNSISVAEHVMMILSLVRYNLPSEHW-ARKGWN 179
121 AGIGSDHVDLSAIDRNVTAEVTCNSISVAEHVMMILSLVRYNLPSEHW-ARKGWN 180
QY 180 IACVSHAYDLEAMHVTAAAGRIGLAVRLAPFDVHLHTDRHLPSEVEKELNITW 239
181 IADVSRAVDLEMDVGLSGRIQAVLRRLAPFDVRLHYSVHRLPKPEVEEELTWM 240
QY 240 ATREDWYPCVDVTLNCPLEPTEHMINDETLEKFRGAVIYVMTAGKLCDRDAVVALE 299
241 PDARSLASSVDVLSITPLPLOTNLPDDELIGAMKRGSTIYVMTARLLVDRDAVVALE 300
QY 300 SGLIAGYAGDVWPEOPAPKDHMPRTMPYNGMTPHISGTTLLTAQARVAAGTRELCEFE 359
301 SGLIAGYAGDVWYPOPPPDHMPRTMPYNGMTPHISGTTLLTAQARVAAGTRELCEFE 360
QY 360 RPIRDEYLYOGGALAGTGAHYS 383
361 RPIRPEYLYDGGGLAGTGAHYS 384

RESULT 2

US-10-224-567-1
Sequence 1, Application US/10224567
Publication No. US20030087763A1
GENERAL INFORMATION:
APPLICANT: Kobayashi, Akio
APPLICANT: Fukushima, Eiichihiro
APPLICANT: Isogai, Akira
TITLE OF INVENTION: Method to Promote Growth of a Plant
FILE REFERENCE: 026350-077
CURRENT APPLICATION NUMBER: US/10/224,567
CURRENT FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: US 09/517,427
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: JP 11-56,776
PRIOR FILING DATE: 1999-03-04
PRIOR APPLICATION NUMBER: JP 2000-36,153
PRIOR FILING DATE: 2000-02-15
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 376
TYPE: PRT
ORGANISM: Rice formate dehydrogenase
FEATURE:
NAME/KEY: VARIANT
LOCATION: 287
OTHER INFORMATION: xaa - Any Amino Acid
US-10-224-567-1

Query Match 39.3%; Score 838.5; DB 15; Length 376;
Best Local Similarity 49.3%; Pred. No. 1.4e-78;
Matches 170; Conservative 53; Mismatches 115; Indels 7; Gaps 2;

QY 31 YPGGQILPPPKAIDTPGQLGSGVSGELRLPYLESNGHTLVTSKDGDPVSFEKRLVD 90
35 YKGEY-----ADKNP-NFVCGVEGALGIREWLESKHXYITDDKEGLNSELEKHIED 87
QY 91 ADVVIOQPPRALPTEBRIAKAKNLALTAGISDHDVLSAIDRNVTAEVTCNSIS 150
88 MHVLTTPPHPAVSAERIKAKNLALLTAGISDHDVLSAIDRNVTAEVTCNSIS 147
QY 151 VAEHVMMILSLVRYNLPSEHWARKGWNIAVDCVSHAYDLEAMHVTAAAGRIGLAVLR 210
148 VAEDELMLRILILRLNLPGLQVYHGEWVAGIATAYADLEKGTGTGAGRIGLAVLR 207
QY 211 LAEDVHLHTDRHLPSEVEKELNITWATREDMTVCVDVTLNCPLEPTEHMINDET 270
208 LKFFNCLLYHDLKIDPELEKIGAKYEEDLDAMLKCDVYIYNTPLTEKTRGMNKR 267
QY 271 LKFRKAGATVNTARCKLDRDAVVALESGRLAGYAGDVWPEOPAPKDHMPRTMPYNGM 330

DB 268 IAKMKGVIIYVNNARGAIMTQAVADACSSGVAGGVDWFPQAPAKGPMRYMNHAM 327
QY 331 TPIHISGTTLLTAQARVAAGTRELCEFEGRINDEYLYOGGALA 375
DB 328 TPIHISGTTIDQLRYAGVADMDLDRYFKGDEFPVONTYVEGOLA 372

RESULT 3

US-09-815-242-12981
Sequence 12981, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.01A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12981
LENGTH: 343
TYPE: PRT
ORGANISM: staphylococcus aureus
US-09-815-242-12981

Query Match 36.6%; Score 781; DB 9; Length 343;
Best Local Similarity 45.6%; Pred. No. 1.2e-72;
Matches 159; Conservative 64; Mismatches 104; Indels 22; Gaps 6;

QY 40 PKAIDTPGQLGSGVSGELRLPYLESNGHTLVTSKDGDPVSFEKRLVDVAVTSOP 99
10 PEAVGQENOLNITKKA-LSKTFLEBERGHEFTIAD-NED-LDKHLPMYIISAPP 65
QY 100 WPAVLTPEBRIAKAKNLALTAGISDHDVLSAIDRNVTAEVTCNSISVAEHVMMI 159
66 YPAVMTREBRIEAPNKLKAITAGVSDHDVLAASEHNGVVEVTSNYSVAEHVMMI 125
QY 160 LSLVRYNLPSEHWARKGWNIAVDCVSHAYDLEAMHVTAAAGRIGLAVRLAPFDVHL 219
126 LILRLYEBGHRSVGEWNLQVGNHAEHLQKRTIGFGFRIQGLVABERLAPFVNTLO 185
QY 220 YTD-----RHRLEPSEVEKELNITWATREDMTVCVDVTLNCPLEPTEHMINDET 274
186 HYPIINCOOHLKLFY-----SFDELVSTSAITIHAPLTPETDLDKDLKRN 235
QY 275 KRGAVYVNTARCKLDRDAVVALESGRLAGYAGDVWPEOPAPKDHMPRTMPYNGMTPHI 334
236 KKHSLVNTARGIYVRDALVEALSEHLQGYAGDVWYQAPADHPHMPRTMPYNGMTPHI 295
QY 335 SGTTLTAQARVAAGTRELCEFEGRINDEYLYOGGALAGTGAHYS 383


```

: GENERAL INFORMATION:
: APPLICANT: OMURA, SATOSHI
: APPLICANT: IKEDA, HARUO
: APPLICANT: ISHIKAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIBA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHIYUKI
: APPLICANT: HATTORI, MASAHIRA
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-262
: CURRENT APPLICATION NUMBER: US/10/156,761
: CURRENT FILING DATE: 2002-05-29
: PRIOR APPLICATION NUMBER: JP 2001-204089
: PRIOR FILING DATE: 2001-05-30
: PRIOR APPLICATION NUMBER: JP 2001-272697
: PRIOR FILING DATE: 2001-08-02
: NUMBER OF SEQ ID NOS: 15109
: SEQ ID NO 15009
: LENGTH: 325
: TYPE: PRT
: ORGANISM: Streptomyces avermitilis
: US-10-156-761-15009

```

```

Query Match          15.9%  Score 340.5; DB 15; Length 325;
Best Local Similarity 31.6%; Pred. No. 8.2e-27;
Matches 94; Conservative 46; Mismatches 146; Indels 11; Gaps 4;

QY 70 TLVTSDDKDPDSYFEEELVDADVVISQPMFAYLPFERIAKAKNLATAGISDHD 129
DB 35 TVVATDDED-----PALREAHVILT-GIGP--VTAHEHTAAPELIGCASGFIYVD 85
QY 130 LQASIDRNVAAY--TYCNSISVAEHVMMILSVNYLPBSHMAKGGNINADCVSHA 187
DB 86 LDAARAGLPCVNCSSGAEQONVAEQTFALMLALQALVPAHATLADADWALPRLQRSI 145
QY 188 YDEAMHVGTVAAGRIGLAVLRRLAPPDVHLHYTRHRLPESVKEKELNTFWHATREDMYP 247
DB 146 TELSGKTLGIVGCHIGEVARRAVADMRIVVGRGVGAEKARLGGARHVLDLRL 205
QY 248 VCDVYTLNCPHLPETHEMINDETLEKLRGAYIVNTARGKLCORDAVARALESGRLGYA 307
DB 206 TADVYTLHAPLTEXTRLHLDADRLALDKPTAFVINTARGALIDODLADLKEKALGAG 265
QY 308 GDWFPQAPAKDHWKRMFNGMTPHISGTTTLQAAVYAAGTRILECFEGRIIRD 364
DB 266 IDVFDPPTPSALRLRLAPNVVLSPHVAGVTRTLVRIALAAVONADFYAGETPRD 322

```

```

RESULT 7
US-09-738-626-4915
: Sequence 4915, Application US/09738626
: Publication No. US20020197605A1
: GENERAL INFORMATION:
: APPLICANT: NAKAGAWA, SATOSHI
: APPLICANT: MIZOGUCHI, HIROSHI
: APPLICANT: ANDO, SEIKO
: APPLICANT: HAYASHI, MIKIRO
: APPLICANT: OCHIAI, KEIKO
: APPLICANT: YOKOI, HARUHIKO
: APPLICANT: TATEISHI, NAOKO
: APPLICANT: SENOH, AKIHITO
: APPLICANT: IKEDA, MASATO
: APPLICANT: OZAKI, AKIO
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-125
: CURRENT APPLICATION NUMBER: US/09/738,626
: CURRENT FILING DATE: 2000-12-18
: PRIOR APPLICATION NUMBER: JP 99/377484
: PRIOR FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03

```

```

: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 4915
: LENGTH: 530
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
: US-09-738-626-4915

```

```

Query Match          14.5%  Score 310.5; DB 10; Length 530;
Best Local Similarity 28.9%; Pred. No. 2.3e-23;
Matches 103; Conservative 51; Mismatches 149; Indels 53; Gaps 12;

QY 64 LESNGHTLVVTSK-----DGPDSYFEEELVD-----ADVVISQPFWP 101
DB 1 MSQNGRPVVLITADKLAOSTVDALGDAVEVWVQDPN--RPPELLDAKENDALLVRS--A 55
QY 102 AYLPFERIAKAKNLKLTAGISDHDLDQASIDRNVTVAEVYTCNSISVAEHVMMILS 161
DB 56 TTYDAEYIAAPNLIKIVGRAGVGIDNDIPATEAGVAVNAPTSNHSACEHAISLLS 115
QY 162 LVNNYLPBSHMAKGGNINADCVSHAYDLAMHVGTVAAGRIGLAVLRRLAPPDVHLHYT 221
DB 116 TARQIPADATFLREGKRRSS--ENGVEIRGKTIVGIFGHIGLQFLAFAETTLVAV 173
QY 222 DRHRLPESVKEKELNTFWHATREDMYPVCDVYTLNCPHLPETHEMINDETLEKLRGAYIV 281
DB 174 DRYANPARA--AQLVVEL--VELDELMSRSDVYTHLPKTKTAGFDAQLLAKSKGQIIL 231
QY 282 NTARGKLCORDAVARALESGRLAGYAGDVWPPQAPDHPMTPYNGMTPHISGTLTA 341
DB 232 NAARGGLYDEQALDALESCHIRAGFDVSTFPC-IDSLPLKPOVVYTPHGLASTEEA 290
QY 342 QARYAAGTRILECFEGRIIRDEYLYVGGALAGTGAHSYSGKNATG--SEEA 395
DB 291 QDR--AGT-----DVADSVLKALAGEFVADAV-----NVSGRNVGEVA 327

```

```

RESULT 8
US-10-195-548-2
: Sequence 2, Application US/10195548
: Publication No. US20030008358A1
: GENERAL INFORMATION:
: APPLICANT: MIKIKO SUGA, Masakazu SUGIMOTO, Tsuyoshi OSUMI, Tsuyoshi
: APPLICANT: NAKAMATSU, Wataru HIBINO, Mika ITO
: TITLE OF INVENTION: METHOD OF PRODUCING L-SERINE BY FERMENTATION
: FILE REFERENCE: OP812
: CURRENT APPLICATION NUMBER: US/10/195,548
: CURRENT FILING DATE: 2002-07-16
: PRIOR APPLICATION NUMBER: US/09/222,786
: PRIOR FILING DATE: 1998-12-30
: PRIOR APPLICATION NUMBER: JP 10-3751
: PRIOR FILING DATE: 1998-01-12
: PRIOR APPLICATION NUMBER: JP 10-353513
: PRIOR FILING DATE: 1998-12-11
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 345
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
: US-10-195-548-2

```

```

Query Match          14.5%  Score 309.5; DB 15; Length 345;
Best Local Similarity 28.9%; Pred. No. 1.5e-23;
Matches 103; Conservative 51; Mismatches 149; Indels 53; Gaps 12;

QY 64 LESNGHTLVVTSK-----DGPDSYFEEELVD-----ADVVISQPFWP 101
DB 1 VSQNGRPVVLITADKLAOSTVDALGDAVEVWVQDPN--RPPELLDAKENDALLVRS--A 55
QY 102 AYLPFERIAKAKNLKLTAGISDHDLDQASIDRNVTVAEVYTCNSISVAEHVMMILS 161
DB 56 TTYDAEYIAAPNLIKIVGRAGVGIDNDIPATEAGVAVNAPTSNHSACEHAISLLS 115

```

RESULT 10

RESULT 11
US-10-156-761-10267
Sequence 10267, Application US/10156761
Publication NO. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDE
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156, 761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02

Matches	89; Conservative	44; Mismatches	117; Indels	30; Gaps	7
QY	76 DKDGDSVERELV-----DADVVISQPFMPAYLTPERIA---KANNLKLALTAGIGSDH	127			

Db 51 EQGEAEVALPEALLQGTDDIDILITQ-----PAPVTAIVEDKIPKIKYGVLRGIVEN 104
Qy 128 VDLQSAIDRNVTAVEVYCNISIAEHVVMILSVRNYPESHENARRGNINADCVSHA 187
Db 105 VNLQVANARGEVNMTPGRNARSVAEFVGMILAEIRNISHDLRDKYR-KDSPHQ 163
Qy 188 Y--PLEMNVGTVAAGRIGLAVRLAPEDVHLHYTD-----HRLPESEVEKELNLTWAT 241
Db 164 AIPELGKRVGLVGLGHTAQVAGFLSGFTEILFYDKYVAGHERYKVD-----S 214
Qy 242 REDMPCVDVYTLCPLEPTEHMINDETLLKFKRGAYIVNTARGKLCRDVAVARALESG 301
Db 215 LDQLVQRADVSIARARLPETENLINAHFALMKRSALTIVTARSGLINEREMIDALNSG 274
Qy 302 RLACGADVWFPPQAPKDPWRTMPYNGMTPHISGTLTJA 341
Db 275 QIMGAAIDTFDEPLPDDSAFYTLNNTVITPHIAGSTIDA 314

RESULT 14

US-10-177-293-92
; Sequence 92, Application US/10177293
; Publication No. US20030124128A1

GENERAL INFORMATION:

APPLICANT: Lillie, James
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Gannavarpu, Manjula
APPLICANT: Kamatkar, Shubhang
APPLICANT: Mertens, Maureen
APPLICANT: Myer, Vic
APPLICANT: Wang, Youzhen
APPLICANT: Xu, Yongyao
APPLICANT: Hoersch, Sebastian
APPLICANT: Monahan, John
APPLICANT: Meyers, Rachel E.
APPLICANT: Bast Jr., Robert C.
APPLICANT: Hortobagyi, Gabriel N.
APPLICANT: Pusztai, Lajos
APPLICANT: Meric, Funda
APPLICANT: Sahin, Aysegul
TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST CANCER
TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-038
CURRENT APPLICATION NUMBER: US/10/177,293
CURRENT FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: US 60/299,887
PRIOR FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: US 60/301,572
PRIOR FILING DATE: 2001-06-27
PRIOR APPLICATION NUMBER: US 60/306,501
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/325,002
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 60/362,585
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/xxx,xxx
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 506
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 92
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens
US-10-177-293-92

Query Match 12.9%; Score 275.5; DB 15; Length 445;
Best Local Similarity 28.0%; Pred. No. 7.9e-20;
Matches 90; Conservative 56; Mismatches 143; Indels 33; Gaps 7;
Qy 59 GLRPLSENG--HTLVVTSKDGPDVSFE-----RELVDADVVI 95

Db 20 GTRPOI-MNGPLHRRPVALLDGRCTYEMPIKDLATVAFCDAGSTOEIHEKVLNEAVG 78
Qy 96 SQPWPAYLTPERIKAKANLKIALTAGIGSDHVDLSQSIDRNVTVAEVYCNISIAEHV 155
Db 79 AMMYHTLTREDLEKFKALRYIVIGSGYDNDIKKAGEIGIAVCNIPSAVEERTADST 138
Qy 156 VMILSVRNRLPSHEARKG-----WNIDCVSHADLEMHNVGTVAAGRIGLAVLR 210
Db 139 ICHINLTNRRTWLYQALREGTRQGVSEQIREVNSGARRIGETGLIGFPGTGOAVAVR 198
Qy 211 LAPFDVHLYTDRHRLPSEVEKELNLTWHTAREDMYPCVDVYTLCPLEPTEHMINDET 270
Db 199 AKAREFSYTFDPT-LQGIERSLGVQVYTLQDLVYSDCVSLHCNINENHHLINPT 257
Qy 271 LKLFKRGAYIVNTARGKLCRDVAVARALESGRLACGADVWFPP-APKDPWRTMPYNG 329
Db 258 IKOMQGAFLVNAAGLVDKALAQALKEGRIGALDVHSEPFSPAGPLKDAPNLI 317
Qy 330 MTPHIS--GTLTIAQARAAGT 349
Db 318 CTPHTAWYSEQASLEMRBAAT 339

RESULT 15

US-10-205-823-76
; Sequence 76, Application US/10205823
; Publication No. US20030108963A1

GENERAL INFORMATION:

APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarpu, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhang
APPLICANT: Monsey, Angela M.
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF PROSTATE CANCER
TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 76
LENGTH: 985
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-76

Query Match 12.4%; Score 264; DB 15; Length 985;
Best Local Similarity 26.7%; Pred. No. 4.3e-18;
Matches 95; Conservative 61; Mismatches 160; Indels 40; Gaps 9;
Qy 32 PGGQLPFP-----KAIDFPQQLIGS--VSGELCLRPYLENSG--HTLVVTSKDGPD 82
Db 526 PASOSLHPHPSPYEXVARRTGAPLIVSYMLAPESIRPOI-MNGPLHRRPVALLDGRDC 584
Qy 83 VFE-----RELVDADVVISQPPWPAYLTPERIRAKKKNLKIALTA 121

Db	585	TVEPRLKDLATVAFCDAQSTQETIHERVLEAVGAMMYHTTENREDLEKFKALRYIVRI	644
QY	122	GIGSDHYDLOSAIDRNTVAEVTYCNSISVAEHVMMILSLVRNYLPSHEWARKG----	177
Db	645	GSGYDNVNKAKAGELGIAVCNIPSAVEETADSTICHILNLYRRNTWLYQALREGTRVQS	704
QY	178	-WNIADCVSHAYDLEAMHVGTVAAGRIGLAVLRRLAPFDVHLHYTDHRRLPESYEKELNL	236
Db	705	VEQIREVAGSAGARIIGETTLGIGFGRGQAVAVRAKAFGSVIFYPDY-LQDGIERSLGV	763
QY	237	TWHTREDMYPVCDVYTLNCPLEHPEHEMINDETLKFKGAVYVNTARGKLCDDAVAR	296
Db	764	QRYTTLODDLXQSDCVSLHCNLEHNNHLLINDEFTIKOMROGAFLVNARGGLVDEKALAQ	823
QY	297	ALBSGRLAGYAGDYWFPQP-APKDPWRTPYNGMTPPHIS--GTTLTAQARYAGT	349
Db	824	AKRGRIKGAALDVHESEPFSAOGPLKDAFNLICHTHTAWYSEQASLEHREAAAT	879

Search completed: July 24, 2003, 20:44:25
 Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:43:38 ; Search time 99 Seconds
(without alignments)
1045.243 Million cell updates/sec

Title: US-09-996-008B-2

Perfect score: 2135
Sequence: 1 MAKVLCVLDDPDVGYPKTY.....YSKGNATGSGSEAKFKKAV 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPREMBL_23:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2127	99.6	401	2	Q93GVL mycobacteri
2	2122	99.4	400	2	Q9R5V0 mycobacteri
3	1869	87.5	402	2	Q08375 moraxella s
4	1806	84.6	400	2	Q93GWL paracoccus
5	1710	80.1	399	2	Q93JW1 hyphomicrob
6	1704	79.8	401	16	Q930E7 thizobium m
7	1511	70.8	398	2	Q9F7P9 uncultured
8	911	42.7	417	3	Q9Y790 mycophaere
9	821	38.5	364	3	Q93968 candida boi
10	818	38.3	364	3	Q13437 candida met
11	812	38.0	364	3	Q00498 candida met
12	803	37.6	376	3	Q08911 saccharomyc
13	781	36.6	374	16	Q8N9N1 staphylococ
14	777	36.4	374	16	Q99X44 staphylococ
15	520	24.4	248	10	Q8VX85 pinus pinas
16	502	23.5	199	10	Q8W520 zea mays (m

17	436	20.4	236	3	Q08988 saccharomyc
18	405.5	19.0	522	17	Q8TYK0 methanopyru
19	373	17.5	326	17	Q9Y8X7 aeropyrum p
20	352	16.5	333	17	Q9UYH9 pyrococcus
21	350.5	16.4	306	17	Q8U135 pyrococcus
22	348.5	16.3	528	16	Q8EN61 oceanobacil
23	347.5	16.3	323	17	Q8ZXX8 pyrobaculum
24	347	16.3	145	3	Q08987 saccharomyc
25	346.5	16.2	333	17	Q58256 pyrococcus
26	346	16.2	333	17	Q8U3T5 pyrococcus
27	345.5	16.2	307	17	Q9Y0M8 pyrococcus
28	340	15.9	354	16	Q8UJZ6 agrobacteri
29	337.5	15.8	335	17	Q9UYR1 pyrococcus
30	336.5	15.8	307	17	Q50095 pyrococcus
31	332	15.6	323	17	Q28495 archaeoglob
32	332	15.6	328	16	Q92YX6 thizobium m
33	326	15.3	376	17	Q58320 pyrococcus
34	323	15.1	449	13	Q9DEG7 brachydanio
35	322	15.1	540	17	Q8PW48 methanosc
36	319	14.9	322	2	Q59463 hyphomicrob
37	316	14.8	316	16	Q9CN80 pasteurella
38	313.5	14.7	534	16	Q8NW42 staphylococ
39	312.5	14.6	540	16	Q9KC99 bacillus ha
40	312	14.6	309	17	Q9HK29 thermoplasm
41	311.5	14.6	534	16	Q99TE0 staphylococ
42	310.5	14.5	530	16	Q8NGY7 coynebacte
43	309.5	14.5	530	16	Q8FPV9 coynebacte
44	308.5	14.4	334	16	Q92T15 thizobium m
45	308	14.4	360	16	Q8YEC6 bruceella me

ALIGNMENTS

RESULT 1	ID	Q93GVL	PRELIMINARY:	PRT:	401 AA.
AC	Q93GVL				
DT	01-DEC-2001	(TREMBLrel. 19, Created)			
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)			
DE	Formate dehydrogenase (EC 1.2.1.2).				
GN	FDH.				
OS	Mycobacterium vaccae.				
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.				
OX	NCBI_TaxID=1810;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;				
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96158475; PubMed=8597552;				
RA	Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;				
RT	"Cloning of formate dehydrogenase gene from a methanol-utilizing				
RT	bacterium Mycobacterium vaccae N10.";				
RT	bacterium Mycobacterium vaccae N10.";				
RT	Appl. Microbiol. Biotechnol. 44:479-483(1995).				
DR	EMBL; AB072394; BAB69476.1; .				
DR	InterPro: IPR006139; 2-Hacid_DH.				
DR	InterPro: IPR006140; 2-Hacid_DH.C.				
DR	Pfam; PF00389; 2-Hacid_DH.1.				
DR	Pfam; PF02826; 2-Hacid_DH.C; 1.				
DR	PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.				
DR	PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.				
DR	PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.				
KW	Oxidoreductase.				
SO	SEQUENCE 401 AA; 44149 MW; 864821B4790AAD26 CRC64;				
Query Match					
Best Local Similarity	99.6%; Score 2127; DB 2; Length 401;				
Matches	400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;				

```

QY 1 MAKVLCVLYDDPVDGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELGL 60
DB 1 MAKVLCVLYDDPVDGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELGL 60
QY 61 RPYLESNGHTLVVTSDDGDPDSVFERELVDADVITISOPFMPAYLTPPERIAKAKNLKTLAT 120
DB 61 RPYLESNGHTLVVTSDDGDPDSVFERELVDADVITISOPFMPAYLTPPERIAKAKNLKTLAT 120
QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPESHMARKGWN 180
DB 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPESHMARKGWN 180
QY 181 ADCVSHAYDLEAMHVGVAAGRGIGLAVLRRLAPFDVHLHTDRHRLEPESVEKELNTLWHA 240
DB 181 ADCVSHAYDLEAMHVGVAAGRGIGLAVLRRLAPFDVHLHTDRHRLEPESVEKELNTLWHA 240
QY 241 TREDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 300
DB 241 TREDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 300
QY 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTRELLECFEGR 360
DB 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTRELLECFEGR 360
QY 361 PIRDEYLIYVGGALAGTGAHSGYSKGNATGSGSEAAKFKKAV 401
DB 361 PIRDEYLIYVGGALAGTGAHSGYSKGNATGSGSEAAKFKKAV 401

```

RESULT 2

Q9RSV0 PRELIMINARY; PRT; 400 AA.

```

ID 09RSV0 PRELIMINARY; PRT; 400 AA.
AC Q9RSV0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE NAD(+)-dependent formate dehydrogenase (EC 1.2.1.2) (Fragment).
OS Mycobacterium vaccae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1810;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96158475; PubMed=8597552;
RA Galkin A., Kulakova L., Tishkov V., Esaki N., Soda K.;
RT "Cloning of formate dehydrogenase gene from a methanol-utilizing
RT bacterium Mycobacterium vaccae N10."
RL Appl. Microbiol. Biotechnol. 44:479-483(1995).
DR HSSP; P33160; 2-NAD.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH.C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
SQ SEQUENCE 400 AA; 44018 MW; 4422888CA22BD9B CRC64;

```

Query Match 99.4%; Score 2122; DB 2; Length 400;

Best Local Similarity 99.8%; Pred. No. 2.9e-166;

Matches 399; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 2 AKVLCVLYDDPVDGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELGL 61
DB 1 AKVLCVLYDDPVDGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELGL 60
QY 62 RPYLESNGHTLVVTSDDGDPDSVFERELVDADVITISOPFMPAYLTPPERIAKAKNLKTLAT 121
DB 62 RPYLESNGHTLVVTSDDGDPDSVFERELVDADVITISOPFMPAYLTPPERIAKAKNLKTLAT 120
QY 122 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPESHMARKGWN 181
DB 122 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPESHMARKGWN 180
QY 181 ADCVSHAYDLEAMHVGVAAGRGIGLAVLRRLAPFDVHLHTDRHRLEPESVEKELNTLWHA 240
DB 181 ADCVSHAYDLEAMHVGVAAGRGIGLAVLRRLAPFDVHLHTDRHRLEPESVEKELNTLWHA 240
QY 241 TREDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 300
DB 241 TREDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 300
QY 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTRELLECFEGR 360
DB 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTRELLECFEGR 360

```

```

QY 182 DCVSHAYDLEAMHVGVAAGRGIGLAVLRRLAPFDVHLHTDRHRLEPESVEKELNTLWHA 241
DB 182 DCVSHAYDLEAMHVGVAAGRGIGLAVLRRLAPFDVHLHTDRHRLEPESVEKELNTLWHA 240
QY 242 REDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 301
DB 242 REDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 300
QY 241 REDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 300
DB 241 REDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 300
QY 302 RLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTRELLECFEGR 361
DB 302 RLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTRELLECFEGR 360
QY 362 IRDEYLIYVGGALAGTGAHSGYSKGNATGSGSEAAKFKKAV 401
DB 362 IRDEYLIYVGGALAGTGAHSGYSKGNATGSGSEAAKFKKAV 400

```

RESULT 3

008375 PRELIMINARY; PRT; 402 AA.

```

ID 008375 PRELIMINARY; PRT; 402 AA.
AC 008375;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE NAD-dependent formate dehydrogenase (EC 1.2.1.2).
GN FDH.
OS Moraxella sp.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Moraxella.
OX NCBI_TaxID=479;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C-1;
RA Galkin A.G.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y13245; CAAT73696.1; -.
DR HSSP; P33160; 2-NAD.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH.C.
DR Pfam; PF00389; 2-Hacid_DH.1.
DR Pfam; PF02826; 2-Hacid_DH.C.1.
DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
KW NAD; Oxidoreductase.
SQ SEQUENCE 402 AA; 44143 MW; C51CE1CE996BBB CRC64;

```

Query Match 87.5%; Score 1869; DB 2; Length 402;

Best Local Similarity 85.7%; Pred. No. 1.9e-145;

Matches 342; Conservative 26; Mismatches 31; Indels 0; Gaps 0;

```

QY 1 MAKVLCVLYDDPVDGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELGL 60
DB 1 MAKVLCVLYDDPVDGPKTYARDDLPRKIDHYPGQILPTPKAIDFTPGQLGSGVSGELGL 60
QY 61 RPYLESNGHTLVVTSDDGDPDSVFERELVDADVITISOPFMPAYLTPPERIAKAKNLKTLAT 120
DB 61 RPYLESNGHTLVVTSDDGDPDSVFERELVDADVITISOPFMPAYLTPPERIAKAKNLKTLAT 120
QY 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPESHMARKGWN 180
DB 121 AGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEHVMMILSLVNNYLPESHMARKGWN 180
QY 181 ADCVSHAYDLEAMHVGVAAGRGIGLAVLRRLAPFDVHLHTDRHRLEPESVEKELNTLWHA 240
DB 181 ADCVSHAYDLEAMHVGVAAGRGIGLAVLRRLAPFDVHLHTDRHRLEPESVEKELNTLWHA 240
QY 241 TREDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 300
DB 241 TREDMYPVCDVVTLLNCPLEHETEMINDETLKLFKRGAYIVNTARGKLCDDAVARALES 300
QY 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTRELLECFEGR 360
DB 301 GRLAGYAGDVWFPQAPKDPHWRTPYNGMTPHISGTTLLTAQARYAAGTRELLECFEGR 360

```


Db 301 GRLAGYAGDWFPPAPRDPHWRTPMHNHGMTPHISGTSLSQTRAACTREILCEYFEGR 360
OY 361 PIRDEYLIYOGGALAGTAHSGNSKGNATGSGSEAAKFKK 399
Db 361 PIRDEYLIYOGGALAGTAHSGNSKGNATGSGSEAAKFKK 399

RESULT 4

OY36W3 PRELIMINARY; PRT; 400 AA.
AC OY36W3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE NAD-dependent formate dehydrogenase.
GN FDH.
OS Paracoccus sp. 12-A.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Paracoccus.
OX NCBI_TaxID=171192;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12-A;
RA Shinoda T., Satoh T., Mineki S., Iida M., Taguchi H.;
RT "Cloning, nucleotide sequence and expression in Escherichia coli of
the gene for formate dehydrogenase of Paracoccus sp. 12-A, a formate-
assimilating bacterium."
RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB071373; BAB64941.1; -
DR InterPro: IPR006139; 2-Hacid_DH.C.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.C; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
SQ SEQUENCE 400 AA; 43757 MW; 13195AED9AC299D8 CRC64;

Query Match 84.6%; Score 1806; DB 2; Length 400;
Best Local Similarity 83.2%; Pred. No. 3e-140;
Matches 333; Conservative 28; Mismatches 39; Indels 0; Gaps 0;

OY 1 MAKVLCVLYDDPVGPKTYARDLPRKIDHPRGQILPTPKAIDFTPEQLLGVSSELGL 60
Db 1 MAKVLCVLYDDPVGPKTYARDLPRKIDHPRGQILPTPKAIDFTPEQLLGVSSELGL 60
OY 61 RPYLESNGHTLVMTSDKGPDSVFERELVDADVVISOPFWAYLTPERIAKAKNLKALT 120
Db 61 RPYLESNGHTLVMTSDKGPDSVFERELVDADVVISOPFWAYLTPERIAKAKNLKALT 120
OY 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPESHMARKGMNI 180
Db 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPESHMARKGMNI 180
OY 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPESHMARKGMNI 180
Db 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPESHMARKGMNI 180
OY 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPEDVNLHYTDRNRLPESVEKELNLTNHA 240
Db 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPEDVNLHYTDRNRLPESVEKELNLTNHA 240
OY 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPEDVNLHYTDRNRLPESVEKELNLTNHA 240
Db 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPEDVNLHYTDRNRLPESVEKELNLTNHA 240
OY 241 TREDMYPCDVVTLNCPRLHETEMHINDETLLKFRGAYIVNTARGKLCDDDAVARALES 300
Db 241 TREDMYPCDVVTLNCPRLHETEMHINDETLLKFRGAYIVNTARGKLCDDDAVARALES 300
OY 241 TREDMYPCDVVTLNCPRLHETEMHINDETLLKFRGAYIVNTARGKLCDDDAVARALES 300
Db 241 TREDMYPCDVVTLNCPRLHETEMHINDETLLKFRGAYIVNTARGKLCDDDAVARALES 300
OY 301 GRLAGYAGDWFPPAPRDPHWRTPMHNHGMTPHISGTSLSQTRAACTREILCEYFEGR 360
Db 301 GRLAGYAGDWFPPAPRDPHWRTPMHNHGMTPHISGTSLSQTRAACTREILCEYFEGR 360
OY 361 PIRDEYLIYOGGALAGTAHSGNSKGNATGSGSEAAKFKK 400
Db 361 PIRDEYLIYOGGALAGTAHSGNSKGNATGSGSEAAKFKK 400

RESULT 5

OY36W1 PRELIMINARY; PRT; 399 AA.
AC OY36W1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE NAD+ dependent formate dehydrogenase.
GN FDH.
OS Rhizobium sp. JCI17.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=142666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCI17;
RA Tanaka Y., Yoshida T., Watanabe K., Mitsuana T.;
RT "Characterization, gene cloning and expression of NAD+ dependent
formate dehydrogenase from a methylophilic bacterium Rhizobium
sp. JCI17."
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB051073; BAB55449.1; -
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR Pfam: PF00389; 2-Hacid_DH.C; 1.
DR Pfam: PF02826; 2-Hacid_DH.C; 1.
DR PROSITE; PS00671; D_2_HYDROXYACID_DH.3; 1.
SQ SEQUENCE 399 AA; 44044 MW; 853BAD2C635864DC CRC64;

Query Match 80.1%; Score 1710; DB 2; Length 399;
Best Local Similarity 78.2%; Pred. No. 2.4e-132;
Matches 312; Conservative 42; Mismatches 45; Indels 0; Gaps 0;

OY 1 MAKVLCVLYDDPVGPKTYARDLPRKIDHPRGQILPTPKAIDFTPEQLLGVSSELGL 60
Db 1 MAKVLCVLYDDPVGPKTYARDLPRKIDHPRGQILPTPKAIDFTPEQLLGVSSELGL 60
OY 61 RPYLESNGHTLVMTSDKGPDSVFERELVDADVVISOPFWAYLTPERIAKAKNLKALT 120
Db 61 RPYLESNGHTLVMTSDKGPDSVFERELVDADVVISOPFWAYLTPERIAKAKNLKALT 120
OY 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPESHMARKGMNI 180
Db 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPESHMARKGMNI 180
OY 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPESHMARKGMNI 180
Db 121 AGISGDHVDLQSAIDRNTVAEVTYCNISVAEHVVMILSLVRNYLPESHMARKGMNI 180
OY 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPEDVNLHYTDRNRLPESVEKELNLTNHA 240
Db 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPEDVNLHYTDRNRLPESVEKELNLTNHA 240
OY 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPEDVNLHYTDRNRLPESVEKELNLTNHA 240
Db 181 ADCVSHAYDLEAMHVGTAAGRIGLAVLRRLAPEDVNLHYTDRNRLPESVEKELNLTNHA 240
OY 241 TREDMYPCDVVTLNCPRLHETEMHINDETLLKFRGAYIVNTARGKLCDDDAVARALES 300
Db 241 TREDMYPCDVVTLNCPRLHETEMHINDETLLKFRGAYIVNTARGKLCDDDAVARALES 300
OY 241 TREDMYPCDVVTLNCPRLHETEMHINDETLLKFRGAYIVNTARGKLCDDDAVARALES 300
Db 241 TREDMYPCDVVTLNCPRLHETEMHINDETLLKFRGAYIVNTARGKLCDDDAVARALES 300
OY 301 GRLAGYAGDWFPPAPRDPHWRTPMHNHGMTPHISGTSLSQTRAACTREILCEYFEGR 360
Db 301 GRLAGYAGDWFPPAPRDPHWRTPMHNHGMTPHISGTSLSQTRAACTREILCEYFEGR 360
OY 361 PIRDEYLIYOGGALAGTAHSGNSKGNATGSGSEAAKFKK 399
Db 361 PIRDEYLIYOGGALAGTAHSGNSKGNATGSGSEAAKFKK 399

RESULT 6

OY30E7 PRELIMINARY; PRT; 401 AA.
ID OY30E7;
AC OY30E7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Probable NAD-dependent formate dehydrogenase.
GN RA0251 OR SMA0478.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

DR HSSP; P33160; 2NAC.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS0065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 DR Oxidoreductase.
 SW SEQUENCE 364 AA; 40370 MW; 1B30982E0D5B778 CRC64;

Query Match 38.3%; Score 818; DB 3; Length 364;
 Best Local Similarity 44.9%; Pred. No. 5.5e-59;
 Matches 170; Conservative 57; Mismatches 114; Indels 38; Gaps 6;

QY 3 KVLCLYDDPYDGYPKTYARDLPKIDHYPEGQILPFPKAIIDFTPGQLLSVSGELGRP 62
 DB 2 KIVLVLYD-----AGKHADE-----EKLYGCTENKLGIAN 32
 QY 63 YLENSGHTLVYTSKDDGDSFERELVDADVVISOPFMPAYLTPERLAKANLKLATAG 122
 DB 33 WLKQGHLLTTSKDEGTSLELDKHIPDADIIITPPHAYITKERLDKAKNLKLVYAG 92
 QY 123 IGSDDHDLQ--SAIDRNTVAEVTYCNISVAEHVMMILSLVNTYLPSEHMARKGMI 180
 DB 93 VGSDDHDLQ--SAIDRNTVAEVTYCNISVAEHVMMILSLVNTYLPSEHMARKGMI 180
 QY 181 ADCVSHAYDLEAHMGVYAGRGIGLAVLRRLAPD-VHLATYDHRRLPESVEKELINTW 239
 DB 153 AAIKADYDIEGKTATIGAGRGIRYLERLPPNPKLLLYDQALPKEAEVVGARV 212
 QY 240 ATRDDMTPVCVYVTLNCPLETHEMINDETLKFKGAVLVNARGKLCDDRAVARALE 299
 DB 213 ENIELVAQADIVYVNPALHAGTGLINKELLSFKKGAWLVNARGKLCDDRAVARALE 272
 QY 300 SGRLAGYAGDVWFPQAPKDPHWPRTM-----PYNGMTPHISGTLTQAARAACTRELE 354
 DB 273 SGQLRGIGGVWFPQAPKDPHWPRTM-----PYNGMTPHISGTLTQAARAACTRELE 354
 QY 355 CFEEGR-PIRDEYLIYOGG 372
 DB 333 SFFGKFDYRPQDITLLNG 351

RESULT 11
 ID 000498 PRELIMINARY; PRT; 364 AA.
 AC 000498;
 DT 01-NOV-1996 (TREMUREL. 01, Created)
 DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
 DT 01-OCT-2002 (TREMUREL. 22, Last annotation update)
 DE NAD-dependent formate dehydrogenase.
 OS Candida methylca.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=45353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96009885; PubMed=7557425;
 RA Allen S.J., Holbrook J.J.;
 RT "Isolation, sequence and overexpression of the gene encoding NAD-
 RT dependent formate dehydrogenase from the methylotrophic yeast Candida
 RT methylca.";
 RL Gene 162:99-104(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hall L.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X81129; CAAS7036.1; -.
 DR HSSP; P33160; 2NAC.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam; PF00389; 2-Hacid_DH; 1.

DR Pfam; PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS0065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 DR NAD.
 SW SEQUENCE 364 AA; 40344 MW; D83DD1FA9D5EA303 CRC64;

Query Match 38.0%; Score 812; DB 3; Length 364;
 Best Local Similarity 44.6%; Pred. No. 1.7e-58;
 Matches 169; Conservative 57; Mismatches 115; Indels 38; Gaps 6;

QY 3 KVLCLYDDPYDGYPKTYARDLPKIDHYPEGQILPFPKAIIDFTPGQLLSVSGELGRP 62
 DB 2 KIVLVLYD-----AGKHADE-----EKLYGCTENKLGIAN 32
 QY 63 YLENSGHTLVYTSKDDGDSFERELVDADVVISOPFMPAYLTPERLAKANLKLATAG 122
 DB 33 WLKQGHLLTTSKDEGTSLELDKHIPDADIIITPPHAYITKERLDKAKNLKLVYAG 92
 QY 123 IGSDDHDLQ--SAIDRNTVAEVTYCNISVAEHVMMILSLVNTYLPSEHMARKGMI 180
 DB 93 VGSDDHDLQ--SAIDRNTVAEVTYCNISVAEHVMMILSLVNTYLPSEHMARKGMI 180
 QY 181 ADCVSHAYDLEAHMGVYAGRGIGLAVLRRLAPD-VHLATYDHRRLPESVEKELINTW 239
 DB 153 AAIKADYDIEGKTATIGAGRGIRYLERLPPNPKLLLYDQALPKEAEVVGARV 212
 QY 240 ATRDDMTPVCVYVTLNCPLETHEMINDETLKFKGAVLVNARGKLCDDRAVARALE 299
 DB 213 ENIELVAQADIVYVNPALHAGTGLINKELLSFKKGAWLVNARGKLCDDRAVARALE 272
 QY 300 SGRLAGYAGDVWFPQAPKDPHWPRTM-----PYNGMTPHISGTLTQAARAACTRELE 354
 DB 273 SGQLRGIGGVWFPQAPKDPHWPRTM-----PYNGMTPHISGTLTQAARAACTRELE 354
 QY 355 CFEEGR-PIRDEYLIYOGG 372
 DB 333 SFFGKFDYRPQDITLLNG 351

RESULT 12
 ID 008911 PRELIMINARY; PRT; 376 AA.
 AC 008911;
 DT 01-NOV-1996 (TREMUREL. 01, Created)
 DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)
 DT 01-OCT-2002 (TREMUREL. 22, Last annotation update)
 DE Chromosome XV reading frame ORF YOR388C.
 GN FDI1 OR YOR388C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Delius H., Hebling U., Hofmann B.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z75296; CAAG9720.1; -.
 DR HSSP; P33160; 2NAD.
 DR SGD; S0005915; FDI1.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam; PF00389; 2-Hacid_DH; 1.
 DR Pfam; PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 SW SEQUENCE 376 AA; 41714 MW; 67ECD46F9DDC2A02 CRC64;

Query Match 37.6%; Score 803; DB 3; Length 376;

Best Local Similarity 44.7%; Pred. No. 9.9e-58;
Matches 181; Conservative 53; Mismatches 115; Indels 56; Gaps 9;

```

QY 3 KVLGCVLDDPVDYGPRTKTYARDLKRIDHYPGGQILPPKALDFTPGQLGSGVSGELRIP 62
DB 5 KVLVLVE-----GKHAHEQE-----KLLGCIENELGIRN 35
QY 63 YLSENGHLYVTSKD--GPDVSPFERELVDADVYISQPPWAPLPRERAKAKNLKALTA 121
DB 36 FIEHQGELVYTTIDKDEPSTVDRELDAEIVITTPPEPAVISNRKLEAPNKLCTYA 95
QY 122 GIGSDHDLQASIDRNVTVAEYTCNSIVAEHVMMILSVRNVLPSHEARKGNMIA 181
DB 96 GVGSDHVDLEAANKRTIVTEVTSNVYSAEHMATILVLRNNGHQOAIENGEMDIA 155
QY 182 DCSHAYDLEAMHGVYAAAGRIGLAVALRLAPD-VHLHYDRHRLPSSEKELNLTWHA 240
DB 156 GVARNEDLEDKIISTVGAGRIGYRLVAFNPKKLLYYDQELPAEALNRLN---EA 212
QY 241 TR-----EDMYPCGVYVTLNCLPHEPEHMINDETLKFKRGAYIVNTAR 285
DB 213 SKLNGKDIYQRYVEKLEDMYQSDVYVITNCPHLKDSKGLFNKKLISHMKDGAIVNPAR 272
QY 286 GKLCDRDAVARALESGRLAGYAGDVWEPQAPAKDHPKRTMPY-----NGMTPHISGTTLT 340
DB 273 GAICVADVDVAEAVKSGRLAGYGGDVWQKQAPAKDHPKRTMNKHQVAMKTVHISGTSLD 332
QY 341 AQARYAAGTRILECFEFGGR-PIRDEYITVGGALACTGANSYSK 384
DB 333 AQRKYAQGVKNILNYSFKKFDYRPODITIVONGSYA--TRAYGOKK 376

```

RESULT 13

```

Q88YNI1 PRELIMINARY; PRT; 374 AA.
AC Q88YNI1;
DT 01-OCT-2002 (TREMblrel. 22, Created)
DT 01-OCT-2002 (TREMblrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE MAD-dependent formate dehydrogenase.
GN FDH OR MW0151.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 358:1819-1827(2002).
DR EMBL; AP004822; BAB94016.1;
DR InterPro; IPR006139; 2-Hacid_DH.C.
DR InterPro; IPR006140; 2-Hacid_DH.C.
DR Pfam; PF00389; 2-Hacid_DH.1;
DR Pfam; PF02826; 2-Hacid_DH.C.1;
DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
KW Complete proteome.
SQ SEQUENCE 374 AA; 41945 MW; 39BE6D84FD58D587 CRC64;

```

Query Match 36.6%; Score 781; DB 16; Length 374;

Best Local Similarity 45.6%; Pred. No. 6.3e-56;
Matches 159; Conservative 64; Mismatches 104; Indels 22; Gaps 6;

```

QY 40 PKAIDFTPGQLGSGVSGELGRPYLESNGHLYVTSDDGPDVSPFERELVDADVYISOPF 99
DB 41 PEAVGQENQLNTRKA-TGLKTFLEEGHGFIIAD-NGED--LDKHLPMDDVITISAPF 96
QY 100 WPAVLTFRIRAKAKNLKALTAGIGSDHVDLQASIDRNVTVAEYTCNSIVAEHVMMI 159
DB 157 LILRNRYEGHQSVGEWMNLQVGNHAEHQHKTIGTFGFRIGQVAERLAPNVTIQ 216

```

```

DB 97 YPAVYTRERIEKAPLKLATITAGVSDHVDLAAASEHNIGVEYTGSTVSAEHAIVMDL 156
QY 160 LSLVANYLPSHEMARKGMNIADCVSHAYDLEAMHGVYAAAGRIGLAVALRLAPDVLH 219
DB 157 LILRNRYEGHQSVGEWMNLQVGNHAEHQHKTIGTFGFRIGQVAERLAPNVTIQ 216
QY 220 YTD-----RHRLPSSEKELNLTWMTATREDMYPVCDVYVTLNCLPHEPEHMINDETLKLE 274
DB 217 HYDPIQDQHLKSV-----SFDELVSTSDATITAPLPEPIDNLEFDKQVLSRM 266
QY 275 KRGAIVYTAGKLCDRDAVARALESGRLAGYAGDVWEPQAPAKDHPKRTMPYNGMTPHI 334
DB 267 KKHSLVYTAGKIVNRALVALASEHLOAGYAGDVWYQAPAPADHPKRTMRNMTATYH 326
QY 335 SGTTLTAQARYAAGTRILECFEFGGRPIRDEYITVGGALACTGANSYSK 383
DB 327 SGMTELAQRIEDGVKIDLEEFNEPPODKDIIIVASGRIA---SKSYT 372

```

RESULT 14

```

Q99X44 PRELIMINARY; PRT; 374 AA.
AC Q99X44;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE MAD-dependent formate dehydrogenase.
GN FDH OR SAV0177 OR SA0171.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohba T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Kanemori M., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanemori M., Matsumaru H., Matsuyama A., Murakami H., Hosoyama A.,
RA Mizutani-U Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hatton M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003358; BAB56339.1;
DR HSSP; P33160; 2NAD.
DR InterPro; IPR006139; 2-Hacid_DH.
DR InterPro; IPR006140; 2-Hacid_DH.
DR Pfam; PF00389; 2-Hacid_DH.1;
DR Pfam; PF02826; 2-Hacid_DH.C.1;
DR PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
DR PROSITE; PS00670; D_2_HYDROXYACID_DH.2; 1.
KW Complete proteome.
SQ SEQUENCE 374 AA; 41931 MW; 4B31FC87DBF2864D CRC64;

```

Query Match 36.4%; Score 777; DB 16; Length 374;

Best Local Similarity 45.3%; Pred. No. 1.4e-55;
Matches 158; Conservative 65; Mismatches 104; Indels 22; Gaps 6;

```

QY 40 PKAIDFTPGQLGSGVSGELGRPYLESNGHLYVTSDDGPDVSPFERELVDADVYISOPF 99
DB 41 PEAVGQENQLNTRKA-TGLKTFLEEGHGFIIAD-NGED--LDKHLPMDDVITISAPF 96
QY 100 WPAVLTFRIRAKAKNLKALTAGIGSDHVDLQASIDRNVTVAEYTCNSIVAEHVMMI 159
DB 97 YPAVYTRERIEKAPLKLATITAGVSDHVDLAAASEHNIGVEYTGSTVSAEHAIVMDL 156
QY 160 LSLVANYLPSHEMARKGMNIADCVSHAYDLEAMHGVYAAAGRIGLAVALRLAPDVLH 219
DB 157 LILRNRYEGHQSVGEWMNLQVGNHAEHQHKTIGTFGFRIGQVAERLAPNVTIQ 216

```

QY 220 YTD-----RRRLPESYEKELNTLWATREDMYPGVDTYLNCPLRPETHEHINDETJLKF 274
 :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 217 HYPIWODHKLSRFV-----SDELTSSSDATTNAPLPETDNLFDXDVLSRM 266
 :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
QY 275 KRCAYIVNTARGKLCDRDVAVALRESGRLAGAGVGWMPPOPARDHPRTWPNGMTPHI 334
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 267 KRKSILVNTRARKIVNRDLAVLAISHLGAGAGVMYPOPARADHPRTIPRNAMTVHY 328
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 335 SCFTLLTAQARYAAGTRELIECFEFGPRIPIDETLIYOGALACTGHNSY 383
 ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 327 SGTGLAQKRILEDGVVDILEREFFNHNEPPRODKDIITYASRIA----SKSTT 372

RESULT 15
Q8VX85

AD	Q8VX85;	PRELIMINARY;	PRT;	248 AA.
AC	08VX85;			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Putative NAD-dependent formate dehydrogenase (EC 1.2.1.2)			
DE	(Fragment)			
OS	Pinus pinaster (Maritime pine).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta			
OC	Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.			
OX	NCBI_TaxID=71647;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Root;			
RA	Dubos C.; Plomton C.;			
RT	"Identification of water-deficit responsive genes in maritime pine			
RT	(Pinus pinaster Ait.) roots using non radioactive cDNA-AFLP.";			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBD databases.			
DR	EMBL; AF300721; CAC83306.1;			
DR	InterPro; IPR006140; 2-Hacid.DH.C.			
DR	Dfam; PF02826; 2-Hacid.DH.C. 1.			
DR	PROSITE; PS00065; D.2_HYDROXYACID_DH_1; 1.			
DR	PROSITE; PS00670; D.2_HYDROXYACID_DH_2; 1.			
DR	PROSITE; PS00671; D.2_HYDROXYACID_DH_3; 1.			
KW	Oxidoreductase; NAD.			
FT	NON_TER	1	248	
FT	NON_TER	1	248	
SQ	SEQUENCE	248 AA;	27370 MW;	45B61CRAFTD55A0C8 CRC64;

Query Match	24.48;	Score 520;	DB 10;	Length 248;
Best Local Similarity	50.88;	Pred. No. 1.1e-34;		
Matches 101; Conservative	29;	Mismatches 69;	Indels 0;	Gaps 0

QY	138	VVAEATYCNSTSVAEHVVMMTLTVRYNLPBSHEARKGGNNADCSYHATDELAHMYGT	197
Db	4	VVAEATYCGNVASVADELMRLTLILMRNFVPOYKOIYEGDWKVAAISYRSDEGKTGT	63
QY	198	VAAGRIGLAVLRRLRFVDVHLHYTDRHRLPESVEKELMTVHAATREDMYPYCDVVTLNCP	257
Db	64	IGAGRIGKELRLKRPFCNCKLLHYDRLSIGPELEKGTATLETTLDENLPRCDVVTINMP	123
QY	258	LHEETHMINDETLLKLFKRGATVVTAGKCLCDRDVAARALESGRLGIAGDWFPOPAR	317
Db	124	LSDRITGMENKERSISKLKGVLIVNANGAALMDADAADASASGIGYSGDWVFPQAP	183
QY	318	KDHPRTPMYNGMTPHISG	336
Db	184	KDHPRSMENHATPHISG	202

Search completed: July 24, 2003, 20:48:32
Job time : 115 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 24, 2003, 20:40:32 ; Search time 24 Seconds

(without alignments)
785.738 Million cell updates/sec

Title: US-09-996-008b-2

Perfect score: 2135
Sequence: 1 NAKVLCLYDPEVDGPKTY.....YSKGNATGGSEPAKFKKAV 401

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2117	99.2	400	1	FDH_PSESR
2	910.5	42.6	375	1	FDH_NEUCR
3	883.5	41.4	377	1	FDH_EMENTI
4	872	40.8	384	1	FDH_ARATH
5	868	40.7	381	1	FDH_SOLTU
6	845.5	39.6	376	1	FDH_ORISA
7	843	39.5	377	1	FDH_ORVU
8	826	38.7	361	1	FDH_PICAN
9	348	16.3	525	1	SERA_METTH
10	338	15.8	525	1	SERA_BACSU
11	323.5	15.2	524	1	SERA_METJA
12	314	14.7	321	1	DHGT_HYPM
13	309.5	14.5	440	1	CTBP_XENLA
14	307.5	14.4	440	1	CTPI_HUMAN
15	306.5	14.4	440	1	CTPI_MOUSE
16	302.5	14.2	430	1	CTPI_RAT
17	302	14.1	437	1	CTBO_XENLA
18	297	13.9	410	1	SERA_HAEIN
19	290	13.6	527	1	SERA_ARCFU
20	287.5	13.5	533	1	SERA_RAT
21	285.5	13.4	533	1	SERA_HUMAN
22	284	13.3	315	1	YF56_HAEIN
23	281.5	13.2	528	1	SERA_MYCLE
24	280	13.1	386	1	CTBP_DRONE
25	277	13.0	313	1	DHGT_METEX
26	276.5	13.0	445	1	CTP2_MOUSE
27	276.5	13.0	624	1	SERA_ARATH
28	275.5	12.9	445	1	CTP2_HUMAN
29	274	12.8	325	1	TKRA_BACSU
30	273.5	12.8	528	1	SERA_MYCTU
31	272	12.7	350	1	YN14_YEAST
32	269.5	12.6	485	1	SERA_MOUSE
33	264	12.4	469	1	SERA_YEAST

34	261	12.2	331	1	DDH_ZYMMO	P30799 zymomonas m
35	258	12.1	409	1	SERA_ECOLI	P08328 escherichia
36	255	11.9	336	1	PTXD_PSEST	O69054 pseudomonas
37	254.5	11.9	466	1	SERA_SCHPO	P87228 schizosacch
38	252	11.8	322	1	VANH_ENTFC	Q05709 enterococci
39	252	11.8	554	1	SERA_SYNG	P73821 synecocyst
40	248	11.6	323	1	VANH_ENTFA	Q47748 enterococcu
41	244	11.4	332	1	LDHD_LACDE	P26297 lactobacill
42	244	11.4	469	1	SE33_YEAST	P40510 saccharomyc
43	241.5	11.3	334	1	YEAA_SCHPO	O14075 schizosacch
44	235	11.0	331	1	DDH_HAEIN	P44501 haemophilus
45	232	10.9	336	1	LDHD_LACHE	P30901 lactobacill

ALIGNMENTS

RESULT 1	FDH_PSESR	STANDARD:	PRT:	400 AA.
AC	P33160;			
DT	01-OCT-1993 (Rel. 27, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate dehydrogenase) (FDH)			
OS	Pseudomonas sp. (strain 101) (Achromobacter parvulus T1).			
OC	Bacteria: Proteobacteria.			
OX	NCBI_TaxID=33067;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92063805; PubMed=1954846;			
RA	Tishkov V.I., Galkin A.G., Egorov A.M.;			
RT	"NAD-dependent formate dehydrogenase of methylotrophic bacteria pseudomonas sp. 101: cloning, expression, and study of the genetic structure.";			
RL	Dokl. Akad. Nauk SSSR 317:745-748 (1991).			
RN	[2]			
RP	SEQUENCE OF 1-193.			
RX	MEDLINE=90290536; PubMed=2357236;			
RA	Popov V.O., Shumilin I.A., Ustunlikova T.B., Lamzin V.S., Egorov T.A.;			
RT	"NAD-dependent formate dehydrogenase from methylotrophic bacteria pseudomonas sp. 101. I. Amino acid sequence.";			
RL	Bioorg. Khim. 16:324-335 (1990).			
RN	[3]			
RP	X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).			
RX	MEDLINE=92283270; PubMed=1597184;			
RA	Lamzin V.S., Aleshin A.E., Strokopylov B.V., Yukhnevich M.G., Popov V.O., Harutyunyan E.H., Wilson K.S.;			
RT	"Crystal structure of NAD-dependent formate dehydrogenase.";			
RL	Eur. J. Biochem. 206:441-452 (1992).			
RN	[4]			
RP	MUTAGENESIS OF CYS-255.			
RX	MEDLINE=93249485; PubMed=8484798;			
RA	Tishkov V.I., Galkin A.G., Marchenko G.N., Egorova O.A., Sheluho D.V., Yakovleva L.B., Dementeva L.A., Egorov A.M.;			
RT	"Catalytic properties and stability of a Pseudomonas sp. 101 formate dehydrogenase mutants containing Cys-255-Ser and Cys-255-Met replacements.";			
RL	Biochem. Biophys. Res. Commun. 192:976-981 (1993).			
CC	-1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.			
CC	-1- SUBUNIT: Homodimer.			
CC	-1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.			
CC	PIR: J00334; J00334.			
DR	PDB: 2NAC; 26-JAN-95.			
DR	PDB: 2NAD; 26-JAN-95.			
DR	InterPro: IPR006139; 2-Hacid_DH.			
DR	InterPro: IPR006140; 2-Hacid_DH.C.			
DR	InterPro: PFO0289; 2-Hacid_DH.			
DR	Pfam: PFO2826; 2-Hacid_DH_C; 1.			
DR	PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.			
DR	PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.			

DR PROSITE; PS00671; D_2-HYDROXYACID_DH_3; 1.
 KW Oxidoreductase; NAD; 3D-structure.
 FT INIT_MET 0 0
 FT DOMAIN 1 146 CATALYTIC.
 FT DOMAIN 147 333 COENZYME-BINDING.
 FT DOMAIN 334 400 CATALYTIC.
 FT NP_BIND 192 226 NAD (ADP PART).
 FT ACT_SITE 5 5 POTENTIAL.
 FT ACT_SITE 284 284 SUBSTRATE BINDING.
 FT MUTAGEN 255 255 C->S, M: HIGH RESISTANCE TO INACTIVATION
 BY HG(2+), INCREASED STABILITY AT 25 C
 AND DECREASED THERMOSTABILITY AT 45 C.
 FT CONFLICT 77 77 D -> S (IN REF. 1).
 FT CONFLICT 138 139 TV -> VT (IN REF. 1).
 FT CONFLICT 145 145 C -> V (IN REF. 1).
 FT CONFLICT 215 216 VH -> HV (IN REF. 1).
 FT CONFLICT 327 327 N -> D (IN REF. 1).
 FT STRAND 2 6
 FT STRAND 8 8
 FT TURN 12 13
 FT TURN 31 32
 FT TURN 46 47
 FT STRAND 52 52
 FT TURN 53 55
 FT HELIX 56 58
 FT HELIX 60 65
 FT TURN 66 67
 FT STRAND 69 73
 FT TURN 79 80
 FT HELIX 82 87
 FT TURN 88 89
 FT STRAND 92 96
 FT TURN 97 98
 FT STRAND 99 99
 FT STRAND 103 103
 FT HELIX 105 110
 FT TURN 112 113
 FT STRAND 116 119
 FT TURN 125 126
 FT STRAND 127 127
 FT HELIX 129 134
 FT TURN 135 136
 FT STRAND 138 141
 FT TURN 143 144
 FT HELIX 147 162
 FT TURN 163 164
 FT HELIX 165 173
 FT TURN 174 175
 FT HELIX 179 183
 FT TURN 184 185
 FT TURN 190 191
 FT STRAND 193 197
 FT HELIX 201 210
 FT HELIX 211 213
 FT TURN 214 214
 FT STRAND 216 220
 FT TURN 227 233
 FT HELIX 234 234
 FT STRAND 235 237
 FT HELIX 241 244
 FT HELIX 245 247
 FT STRAND 250 253
 FT TURN 259 263
 FT STRAND 265 265
 FT HELIX 267 270
 FT TURN 271 272
 FT TURN 275 276
 FT STRAND 277 281
 FT HELIX 285 287
 FT STRAND 288 288
 FT HELIX 290 298
 FT TURN 299 300
 FT STRAND 302 307

FT TURN 317 318
 FT HELIX 320 322
 FT TURN 323 323
 FT TURN 325 326
 FT STRAND 327 327
 FT HELIX 334 336
 FT HELIX 338 357
 FT TURN 358 358
 FT HELIX 363 365
 FT STRAND 366 368
 FT STRAND 373 373
 FT TURN 375 375
 FT HELIX 376 380
 SQ SEQUENCE 400 AA; 44005 MM; A354151D4312991F CRC64;
 Query Match 99.28; Score 2117; DB 1; Length 400;
 Best Local Similarity 99.58; Pred. No. 9, 1e-172;
 Matches 398; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 AKVLCVLYDDPVDPGPKTYARDLPLKIDHYPGGQILPTPKAIDFTPGQLSGVSGELGLR 61
 DB 1 AKVLCVLYDDPVDPGPKTYARDLPLKIDHYPGGQILPTPKAIDFTPGQLSGVSGELGLR 60
 QY 62 PYLESNGHTLVYTSDKGPDSPVFERELVDADVISOFPWPAYLTPERIAKAKNLKALTA 121
 DB 61 KYLESNGHTLVYTSDKGPDSPVFERELVDADVISOFPWPAYLTPERIAKAKNLKALTA 120
 QY 122 GIGSDHYDLSAIDRNTVAEVTYCNISVAEHVMMILSVRYNLPESHMARKGNIA 181
 DB 121 GIGSDHYDLSAIDRNTVAEVTYCNISVAEHVMMILSVRYNLPESHMARKGNIA 180
 QY 182 DCVSHAYDLEAMHVTVAAGRIGLAVLRRLAPFDVHLHYTDNRHLPESVEKEMLTWHT 241
 DB 181 DCVSHAYDLEAMHVTVAAGRIGLAVLRRLAPFDVHLHYTDNRHLPESVEKEMLTWHT 240
 QY 242 REDMYPCDDVYTLNCPHPETEHMINDETLLKLFKRGAYINTARGKLCDDAVARALESG 301
 DB 241 REDMYPCDDVYTLNCPHPETEHMINDETLLKLFKRGAYINTARGKLCDDAVARALESG 300
 QY 302 RLAGYAGDVNFPQAPRDHPKRTMPYNGMTPHISGTYLTQAARAAAGTRETLECFEGGRP 361
 DB 301 RLAGYAGDVNFPQAPRDHPKRTMPYNGMTPHISGTYLTQAARAAAGTRETLECFEGGRP 360
 QY 362 IRDEYLIYOGGALAGTGAHSTYSGNATGSGESEAERKAV 401
 DB 361 IRDEYLIYOGGALAGTGAHSTYSGNATGSGESEAERKAV 400
 RESULT 2
 FDH_NEUCR STANDARD: PRT; 375 AA.
 ID FDH_NEUCR 007103;
 AC 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
 dehydrogenase) (FDH).
 DE FDH OR 99H12.160.
 GN Neurospora crassa.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OC NCB1_TaxID=5141;
 OX [1]
 SEQUENCE FROM N.A.
 STRAIN-74-OR23-1A / FGSC 987;
 RC MEDLINE-93285982; PubMed-8509325;
 RA Chow C.M., Rajbhandary U.L.;
 RT "Developmental regulation of the gene for formate dehydrogenase in
 Neurospora crassa."
 RL J. Bacteriol. 175:3703-3709(1993).
 RP [2]
 RC SEQUENCE FROM N.A.
 STRAIN-74-OR23-1A / FGSC 987;

PubMed-12655011;
 RA Mannhaupt G., Montrone C., Haase D., Meves H.-W., Algn V.,
 RA Hohnselt J.D., Partmann B., Nyakatura G., Kempfen F., Mäler J.,
 RA Schulte U.;
 RT "What's in the genome of a filamentous fungus? Analysis of the
 RT Neurospora genome sequence."
 RL Nucleic Acids Res. 31:1944-1954(2003).
 CC -1 CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1 SUBUNIT: Homodimer (By similarity).
 CC -1 DEVELOPMENTAL STAGE: DEVELOPMENTALLY REGULATED. EXPRESSED ONLY
 CC DURING CONIDATION AND EARLY GERMINATION.
 CC -1 SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG. TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).

CC EMBL: L13864; AA99900.1; -
 CC EMBL: AL451018; CAC18252.1; -
 CC PIR: A47117; A47117.
 CC HSSP: P33160; 2NAC.
 CC InterPro: IPR006139; 2-Hacid_DH.
 CC InterPro: IPR006140; 2-Hacid_DH.C.
 CC Pfam: PF00389; 2-Hacid_DH.1.
 CC Pfam: PF02826; 2-Hacid_DH.C.1.
 CC PROSITE: PS00065; D_2_HYDROXYACID_DH.1; 1.
 CC PROSITE: PS00670; D_2_HYDROXYACID_DH.2; 1.
 CC PROSITE: PS00671; D_2_HYDROXYACID_DH.3; 1.
 CC Oxidoreductase: NAD.
 CC NP_BIND: 166 201 NAD (ADP PART) (BY SIMILARITY).
 CC FT ACT_SITE: 259 259 SUBSTRATE BINDING (BY SIMILARITY).
 CC FT DOMAIN: 364 374 ALA-RICH.
 CC SQ SEQUENCE 375 AA; 40956 MW; 3073CB95FB204968 CRC64;

Query Match 42.6%; Score 910.5; DB 1; Length 375;
 Best Local Similarity 48.4%; Pred. No. 1,4e-69;

Matches 196; Conservative 56; Mismatches 112; Indels 41; Gaps 8;

OY 1 MAKVLCVLPDVPDYPKTVARDDPKIDHYFGGQILPTPKAIDPTPGOLGSGVSGEL 60
 DB 1 MAKVLAAYLDGKRGK-----EYF-----ELLTIONELGL 31
 OY 61 RPYLESNGHTLVTSDDKDPVSFERELVDADVISOFPWPAVYLPERRIAKNNKLALT 120
 DB 32 RKMLEDQGHITLVTCDDKGENSTFDEKLEDAEIIITTFPHGYLAERLARKKLLKLV 91
 OY 121 AGISGDHVDLOSADRN--VVAEYTYCNSISVAEHVMMILSVRYNLPSEHAKRGKW 178
 DB 92 AGISGDHVDLNAANKNTNGITVAEYTGNSVVAHVMTLLVLRNVPABEQIOEGRW 151
 OY 179 NIADCVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDV-HLYYDRHRLPRESVEKELNLT 237
 DB 212 RVADLEEMLAOCDDVTINCPLHEKTOGLFKNELISKMKKSWLVNTARGAIVKRDVA 271
 OY 238 WHATREDMYPVCDVVTYVLCPLHPETEMHINDETLLKFRGAVYVNTARGKLCDDRAVARA 297
 DB 152 DVAEAKNEFDLEGVVTGVRIGERYLRKLPDFCKELLYIYQPLSAKEKEIGCR 211
 OY 298 LESGLAGYAGDVMPROPAPKDPWR--TMPY---NGMTPHSGTTLTAQARYAGTKEI 352
 DB 272 LKSGLRKGYGGDVMPROPAPDPHRLRYAKNPFGGNAVPHMSGISLDAOKRYAAGTAYI 331
 OY 353 LECFEFGR-PIRDEYLIYOGGALAGTGAHSYSGKNAATGSGSEPAK 396
 DB 332 IESYLSGKHDRPREDLIYGGDYA---TKSYGERERAAAAAAAK 373

RESULT 3

FDH_EMENTI
 ID FDH_EMENTI STANDARD: PRT: 377 AA.
 AC 003134;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable formate dehydrogenase (EC 1.2.1.2) (NAD-dependent formate
 DE dehydrogenase) (FDH).
 GN ACIA.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=162425;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93101140; PubMed=1465107;
 RA Salscha J.A., Cobbett C.S., Hynes M.J.;
 RT "Characterization of the anda-regulated acia gene of Aspergillus
 RT nidulans".
 RL Mol. Gen. Genet. 235:349-358(1992).
 CC -1 CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1 SUBUNIT: Homodimer (By similarity).
 CC -1 INDUCTION: ACETATE INDUCTION MEDIATED BY ANDA REGULATORY GENE.
 CC -1 SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG. TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).

CC EMBL: Z11612; CAA77687.1; -
 CC PIR: S30088; S30088.
 CC HSSP: P33160; 2NAC.
 CC InterPro: IPR006139; 2-Hacid_DH.
 CC InterPro: IPR006140; 2-Hacid_DH.C.
 CC Pfam: PF00389; 2-Hacid_DH.1.
 CC Pfam: PF02826; 2-Hacid_DH.C.1.
 CC PROSITE: PS00065; D_2_HYDROXYACID_DH.1; 1.
 CC PROSITE: PS00670; D_2_HYDROXYACID_DH.2; 1.
 CC PROSITE: PS00671; D_2_HYDROXYACID_DH.3; FALSE_NEG.
 CC Oxidoreductase: NAD.
 CC NP_BIND: 161 191 NAD (ADP PART) (BY SIMILARITY).
 CC FT ACT_SITE: 271 271 SUBSTRATE BINDING (BY SIMILARITY).
 CC SQ SEQUENCE 377 AA; 41531 MW; 5219F5FEBCC931CE CRC64;

Query Match 41.4%; Score 883.5; DB 1; Length 377;

Best Local Similarity 52.1%; Pred. No. 2.7e-67;

Matches 185; Conservative 48; Mismatches 95; Indels 27; Gaps 6;

OY 47 PGOLGYSVSGELRPLYESNGHTLVTSDDKDPVSFERELVDADVISOFPWPAVYLP 106
 DB 14 PG-LTGTEENELGIRKIMKEBOGHITLVTSDDKGENSTFDEKLEDAEIIITTFPHGYLTA 72
 OY 107 ERIAKNNKLALTGIGSDHVDLOSADRN--VVAEYTYCNSISVAEHVMMILSVR 164
 DB 73 ERLAKNNKLALTGIGSDHVDLNAANKNTNGITVAEYTGNSVVAHVMTLLVLR 132
 OY 165 NYLPSEHAKRGWNIADCVSHAYDLEAMHGTVAAGRIGLAVLRRLAPFDV-HLYYDR 223
 DB 133 NFVPAHDQIRNDMVAANAARKEFDLENKVTGVRIGERYLRKLPDFCKELLYYDY 192
 OY 224 HRLPRESVEKELNLTVAHATREDMYPVCDVVTYVLCPLHPETEMHINDETLLKFRGAVYVNTARGKLCDDRAVARA 297
 DB 193 QPLREVEKEIGARVDSLEEMVSOCDVVTINCPLHEKTRGLFKNELISKMKKPKSALLY 252
 OY 273 -----LKKRAYIYNTARGKLCDDRAVARALESGLAGYAGDVMPROPAPKDPWRMP 326
 DB 253 LIPMLMYHKGSLVNTARGAIVKEDVAELKSGHLRKYGGDVMPROPAPKDPWRMP 312

Qy 327 Y-----NGMTPHISGTLTQAQARYAGTRELIECFEGR-PIRDEYLIWGGAL 375
 Db 313 HPMGGNATVPHMSGTSLAQIRYANGTKALIDYSFGSFPYQODILVHGDDYA 367

RESULT 4

ID	FDH_ARATH	STANDARD	PRT	384 AA.
AC	Q957E4			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-dependent formate dehydrogenase) (FDH).			
GN	FDH1 OR FDH OR AF5614780 OR T9L3_80.			
OS	Arabidopsis thaliana (mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; Eucosids II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	SEQUENCE FROM N.A.			
RA	Nishikawa T., Fukusaki E., Kobayashi A.;			
RT	"Formate dehydrogenase cDNA from Arabidopsis thaliana."			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
RN	SEQUENCE FROM N.A.			
RA	Li R., Bonham-Smith P.C., King J.;			
RT	"Arabidopsis thaliana mRNA for NAD-dependent formate dehydrogenase 1."			
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.			
RN	SEQUENCE FROM N.A.			
RA	Markwell J., Osterman J.C., Olson B.J., Skavdahl M., Ramberg H., Hermann M.C.;			
RT	"Induction of leaf formate dehydrogenase by one-carbon metabolites."			
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.			
RN	SEQUENCE FROM N.A.			
RA	STRAIN=cv. Columbia;			
RT	MEUJINE=21016721; PubMed=11130714;			
RA	Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E., Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K., Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S., Nakazaki N., Natsu K., Okumura S., Shinpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M., Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R., Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J., Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M., Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M., Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P., Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Stromatt C., Wagner-McPherson C., Wolam A., Yeakum M., Bell M., Dedina N., Parnell L., Shah R., Rodriguez M., Hoon See L., Vill D., Baker J., Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M., Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I., Voicikert G., Wambutt R., Duesterhoft A., Stekema W., Pohl T., Entian K.-D., Terry N., Hartley N., Bent E., Johnson S., Langham S.-A., McCullagh B., Robben J., Grymoprez B., Zimmermann W., Ransperger U., Medler H., Balke K., Medler E., Peters S., van Staveren M., Dirks W., Moolman P., Klein lankhorst R., Weltjensger T., Bothe G., Rose M., Haut J., Bernerster S., Hempel S., Feldpuschen M., Lamberth S., Villarroel R., Giesen J., Adiles W., Berts O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H., Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;			

"Sequence and analysis of chromosome 5 of the plant Arabidopsis thaliana."

RT Nature 408:823-826(2000).

RL [3]

RN SEQUENCE FROM N.A.

RA STRAIN=cv. Columbia;

RC SHINOZAKI K., Davis R.W., Ecker J.R., Theologis A.;

RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the

RT SSP consortium (Salk/Stanford/PGEC).";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.

CC -1- SUBCELLULAR LOCATION: Mitochondrial (by similarity).

CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

DR EMBL; AB023897; BAA88683.1; -

DR EMBL; AF208029; AAF19435.1; -

DR EMBL; AF208029; AAF19436.1; -

DR EMBL; AF217195; AAF67100.1; -

DR EMBL; AL391149; CAC01877.1; -

DR EMBL; AY054285; AAL06944.1; -

DR EMBL; AY039609; AAK62664.1; -

DR EMBL; AY081734; AAL87387.1; -

DR PIR; T51423; T51423.

DR HSSP; P33160; 2NAD.

DR InterPro; IPR006139; 2-Hacid_DH.

DR InterPro; IPR006140; 2-Hacid_DH_C.

DR Pfam; PF00389; 2-Hacid_DH; 1.

DR Pfam; PF02826; 2-Hacid_DH_C; 1.

DR PROSITE; PS0065; D_2-HYDROXYACID_DH_1; 1.

DR PROSITE; PS00670; D_2-HYDROXYACID_DH_2; 1.

DR PROSITE; PS00671; D_2-HYDROXYACID_DH_3; FALSE NEG.

DR Oxidoreductase; NAD; Mitochondrion; Transit peptide.

DR TRANSIT 1 27

FT CHAIN 28 384

FT NP_BIND 198 232

FT ACT_SITE 290 290

FT SEQUENCE 384 AA; 42409 MW; A12BA423019D862B CXC64;

SO

Query Match 40.8%; Score 872; DB 1; Length 384;

Best Local Similarity 52.9%; Pred. No. 2.6e-66;

Matches 172; Conservative 44; Mismatches 109; Indels 0; Gaps 0;

Qy 51 LGSYSGELGRIPRYLESNGHTLVTSXKDPDSYERELVDADVISOPEFAYITPERIA 110

Db 56 LGCVENMLGIRDMVDSOGHOYIVYDKEGPDCELEKHIPDLHYIISTFPHAYVTAERIK 115

Qy 111 KAKMLKLTATGAGTSDHVDLOSALDRNVTVAEVTYCNISVAEHVMMIISLVNYP 170

Db 116 KAKMLKLTATGAGTSDHVDLOSALDRNVTVAEVTYCNISVAEHVMMIISLVNYP 175

Qy 171 EMARKGMNADVCVSHAYDLEAMHVTVAAGRIGLAVIRLAPDVHLHYTDRIHPE 230

Db 176 NQVYKGGENVAGIYRAYVDLEKRTIGTVAGRIGKLLQRLKPGCNILVHDLQVAP 235

Qy 231 EKEINLTWHTREDMYPVCDVYVTLNCPHPEMIMDETKLKKRAAYVNTARKGLCD 290

Db 236 EKEGKAFVEDLNMKPCDKDIYVIMPLETETRMFMKELGKLLKCVLVVNNARGAIME 295

Qy 291 RDAVARLESGRLAGVGDWFEPOPAPKDHWRMPYNGMTPHISGTLTQAQARYAG 350

Db 296 RQAVVDVAVESGHIGISGDWDPAPAPKDHWRMPYNGMTPHISGTLTQAQARYAG 355

Qy 351 EILIECFEGRPIRDEYLIWGGAL 375

Db 356 DMLERYFKGEDFPENTENIVADGELA 380

RESULT 5

FDH_SOLTU	STANDARD	PRT	381 AA.
ID	FDH_SOLTU		
AC	007511; 09Zr28;		
DT	01-FEB-1995 (Rel. 31, Created)		

DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
 DE dependent formate dehydrogenase) (FDH).
 GN FDH.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Solanum.
 OC NCBI_TaxId=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. BF15;
 RA Houtton-Cabassa C., Ambard-Bretteville F., Remy R.,
 RA Colas des Francs-Small C.;
 RT "Evidence for multiple copies of formate dehydrogenase genes in
 RT plants: isolation of three potato fdh genes fdh1, fdh2 and fdh3.";
 RL (in) Plant Gene Register PGR98-102.
 RN [2]
 RP SEQUENCE OF 3-381 FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=cv. BF15; TISSUE=Tuber;
 RX MEDLINE=94105343; PubMed=8278546;
 RA Colas des Francs-Small C., Ambard-Bretteville F., Small I.D.,
 RA Remy R.,
 RT "Identification of a major soluble protein in mitochondria from
 RT nonphotosynthetic tissues as NAD-dependent formate dehydrogenase.";
 RL Plant Physiol. 102:1171-1177(1993).
 RN [3]
 RP REVISIONS TO N-TERMINUS.
 RC STRAIN=cv. BF15; TISSUE=Tuber;
 RA Colas des Francs-Small C.C.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 26-54.
 RC STRAIN=cv. BF15; TISSUE=Tuber;
 RA Colas des Francs-Small C., Ambard-Bretteville F., Darpas A.,
 RA Sallantin M., Huet J.-C., Pernollet J.-C., Remy R.,
 RT "Variation of the polypeptide composition of mitochondria isolated
 RT from different potato tissues.";
 RL Plant Physiol. 98:273-278(1992).
 RN [5]
 RP CHARACTERIZATION.
 RX Pubmed=9490763;
 RA Houtton-Cabassa C., Ambard-Bretteville F., Moreau F.,
 RA Davy de Virville J., Remy R., Colas des Francs-Small C.;
 RT "Stress induction of mitochondrial formate dehydrogenase in potato
 RT leaves.";
 RL Plant Physiol. 116:627-635(1998).
 CC -1- FUNCTION: Involved in formate-dependent oxygen uptake coupled to
 CC ATP synthesis.
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- TISSUE SPECIFICITY: Found at high levels in developing tubers, at
 CC intermediate level in stems, veins, stolons, and stamens,
 CC and at low level in leaves and roots.
 CC -1- INDUCTION: Induced very rapidly by wounding, and slower by
 CC darkness, chilling, drought, hypoxia, and treatments with formate,
 CC abscisic acid, serine, sarcosine, pyruvate, acetate, ethanol or
 CC methanol.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
 CC -1- CAUTION: There are two other putative pseudogenes, fdh2 and fdh3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 299991; CABI7080.1; -

DR EMBL; 299992; -; NOT_ANNOTATED_CDS.
 DR EMBL; Z21493; CAA9702.2; -
 DR PIR; J02272; J02272.
 DR HSP; P33160; 2NAD.
 DR InterPro; IPR006139; 2-Hacid_DH.
 DR InterPro; IPR006140; 2-Hacid_DH_C.
 DR Pfam; PF00389; 2-Hacid_DH; 1.
 DR Pfam; PF02826; 2-Hacid_DH_C; 1.
 DR PROSITE; PS00065; D_2-HYDROXYACID_DH_1; 1.
 DR PROSITE; PS00670; D_2-HYDROXYACID_DH_2; 1.
 DR PROSITE; PS00671; D_2-HYDROXYACID_DH_3; 1.
 KW Oxidoreductase; NAD; Mitochondrion; Transil peptide.
 FT TRANSIT 1
 FT CHAIN 26
 FT NP_BIND 195 229
 FT ACT_SITE 287 287
 SQ SEQUENCE 381 AA: 42038 MW: 9200351AC5A3A00E CRC64:
 Query Match 40.7%; Score 868; DB 1; Length 381;
 Best Local Similarity 49.7%; Pred. No. 5; 6e-66;
 Matches 175; Conservative 48; Mismatches 117; Indels 12; Gaps. 2;
 QY 35 QILPPIPKAI-----DFTPGGLGVSGLRPRYLSNGHTLVYTSKDPDSV 83
 DB 27 QASPGPKIVGYFYRANEYAEKMP-NFEGCAENALGIRWLESKGHOYVTPDKGEPDE 85
 QY 84 FERELVDADVTSOPWPAYLTPERTAKKNLKLATAGIGSDHVDLQSAIDRNTVAEV 143
 DB 86 LEKHIPDLHLVISTFHPFVATVETKRNKLNQLLTLAGISDHDLKAAGLVAVY 145
 QY 144 TYCNISVAEHVVMMLSLVRNYLPSSHEWARKGNINADCVSHAVDLEAMHGVAAAGRI 203
 DB 146 TGSNTVSVAEDELMLRLILVRNLPQHGVINGENMVAIAIRAVDLEKTVGTAGAGRI 205
 QY 204 GLAVLRRLAPEDVHLHYTDRLRPESVEKELMTWHATREDMTPCVDVYTLCPHPETE 263
 DB 206 GILLQRLKPKFCNLIYHRLKMDSELEQIAKFEEDDKMLSKCDIVIVITPLETKR 265
 QY 264 RHINDETLEKFGAYIVYVARGKLCDDRAVVAIALESGLAGYAGDWPPOPAPDHPMR 323
 DB 266 GHPDKERIKAKIKKGVILVNNAGAIMDQAVVDACNSGHIAGISGDWVYTPQAPKDPHR 325
 QY 324 TWPYNGMPPHISGTTLTQAARYAAGTRELTECFEFGRPTRDEYLVYOGGALA 375
 DB 326 YNPQAMPFHISGTTIDAQLRYAAGTKMDLVYFKEGDEPPANVYIKODELA 377
 RESULT 6
 FDH_ORYSA STANDARD; PRT; 376 AA.
 AC Q95XP2;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Formate dehydrogenase, mitochondrial precursor (EC 1.2.1.2) (NAD-
 DE dependent formate dehydrogenase) (FDH).
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriaralioideae; Oryzaceae; Oryza.
 OC NCBI_TaxId=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shiraishi T., Fukusaki E., Kobayashi A.;
 RT "NAD-dependent formate dehydrogenase.";
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG, TO OTHER FDH.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

OX NCBI_TaxID=4905;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Holtenberg C.P., Janowicz Z.;
 RT "DNA molecules coding for FMDH control regions and structured gene for
 RT a protein having FMDH-activity and their uses.";
 RL Patent number EP0299108, 18-JAN-1989.
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY. STRONG, NO OTHER FDI.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: A06214; CA00531.1; -
 DR HSSP: P33160; 2NAD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 DR Oxidoreductase; NAD.
 KW INT_MET 0 PROBABLE.
 FT NP_BIND 164 199 NAD (ADP PART) (BY SIMILARITY).
 FT ACT_SITE 257 257 SUBSTRATE BINDING (BY SIMILARITY).
 SQ SEQUENCE 361 AA; 39779 MM; 0RC001366F9E47B CMC64;

Query Match 38.7%; Score 826; DB 1; Length 361;
 Best Local Similarity 45.1%; Pred. No. 1.9e-62;
 Matches 171; Conservative 63; Mismatches 107; Indels 38; Gaps 6;

OY 3 KVLCLVLDPPDYKPKTYARDLPKIDHPGGLPPPKAIDFPQGLSGVSELDGP 62
 DB 1 KVLCLVLD-----AGKHADE-----ERLYGTENALGIRD 31
 OY 63 YLESNGHTLVTSKDPDSYFERELVDADVVISQPMFAYLPERTAKAKLTLATAG 122
 DB 32 WLEKGGHDVYVTSKESQNSVLEKNIDADVIISTPFRATITERRIDKAKLLLVYAG 91
 OY 123 IGSQHDVLDQ--SAIDRNVTVAEVTYCNSSISVAEHVVMILSLVRYNLPSEHMARKGNI 180
 DB 92 VGSQHDIDLDIYINOSGRDISVLETVGSNVSVAEHVVMMLVLVRNFVPAHQIISGGMV 151
 OY 181 ADCVSHAYDLEAMVGVVAAAGRICGLAVLRRLAPD-VHLHTDNRRLPESYEKELNLTWH 239
 DB 152 AELAKDSFDLEKGIATIGAGRIYRVLERLVAAPRELLYYDQSLKEEKEVGARV 211
 OY 240 ATREDMYPVCDVYVTLNCPLEHETEMINDETLEKFGAYIVNTARGKLCRDVAARLE 299
 DB 212 HDIELEVAQADIYVINGPLHAGSGLVNAELKFKKAGMLVNTARGICVAEYVAAVAK 271
 OY 300 SGRLAGYAGDVWFPQAPKDKHPARTMPY-----NGMPTISGTLTAQARYAAGTRELLE 354
 DB 272 SGLRGYGGGVWFPQAPKDKHPMSMANKYAGANAMPYHSGSVIDAQVRYAAGTKNILE 331
 OY 355 CFEQGR-PINDEYIVOGG 372
 DB 332 SFTQKEDYRPDDILLNG 350

RESULT 9
 ID SERA_METTH STANDARD; PRT; 525 AA.
 AC 027051;
 DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DB D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (PGDH).
 GN SERA OR KMH970.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert R.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Potluer B., Olu D.,
 RA Spadafora R., Vlcare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jivan N., Caruso A., Bush D., Safer H., Patwell D., Pradhanar S.,
 RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 RT delta: functional analysis and comparative genomics.";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) -> 3-
 CC phosphohydroxypyruvate + NADH.
 CC -1- PATHWAY: Serine biosynthesis; first step.
 CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
 CC DEHYDROGENASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000870; AAB85466.1; -
 DR PIR: H69229; H69229.
 DR HSSP: P08328; 1PSD.
 DR InterPro: IPR006139; 2-Hacid_DH.
 DR InterPro: IPR006140; 2-Hacid_DH.C.
 DR InterPro: IPR002912; ACT.
 DR InterPro: IPR006236; PGDH.
 DR Pfam: PF00389; 2-Hacid_DH; 1.
 DR Pfam: PF02826; 2-Hacid_DH.C; 1.
 DR Pfam: PF01842; ACT; 1.
 DR TIGRFAMs: TIGR01327; PGDH; 1.
 DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
 DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; 1.
 DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; 1.
 KW Serine biosynthesis; Oxidoreductase; NAD: Complete proteome.
 FT ACT_SITE 232 232 SUBSTRATE BINDING (BY SIMILARITY).
 FT ACT_SITE 261 261 BY SIMILARITY.
 FT ACT_SITE 279 279 BY SIMILARITY.
 SQ SEQUENCE 525 AA; 56990 MM; DDD63FCFF812B1C5 CMC64;

Query Match 16.3%; Score 348; DB 1; Length 525;
 Best Local Similarity 31.7%; Pred. No. 9.9e-22;
 Matches 97; Conservative 54; Mismatches 131; Indels 24; Gaps 6;

OY 72 VVTSKDKGPDVSFRELVDADVVISQPMFAYL-----TPPRIAKAK 113
 DB 6 VLINDSTINEKGISLEEV-AEYVYVNTITPELLDAIKDPALVYRSTKTYREVIEAP 64
 OY 114 NLKLATAGIGSDHVDLSAIDRNVTVAEVTYCNSSISVAEHVVMILSLVRYNLPSEHMA 173
 DB 65 RLKTIIRAGVGVNDVVAARDRCIMVINAPESTISIVAEHSIGIMLALARKIAIDRSV 124
 OY 174 RKGGMNADCVSHAYDLEAMVGVVAAAGRICGLAVLRRLAPDVLHHTDNRRLPESYEKE 233
 DB 125 KEKWEKNRPM--GIELNGKTLGITGMRIGSQVYVRKAPAGMDIMVYDPISEAAE-E 181
 OY 234 LNTFWHATREDMYPVCDVYVTLNCPLEHETEMINDETLEKFGAYIVNTARGKLCRDA 293

```

Db      182 MGVT-VTDLTETLLRESIDIVIHVLPETRHLLISEDFELMKKQAFIVNARGGIDEDA 240
Oy      294 VARALESGRAGYAGVWFPAPKDHMPRTMPYNGMTPHISGTLTAAQARYAGTREL 353
Db      241 LYRLAKGETAGALADVFEER-PEGSPLELEENVLTFPHIGASTSEADAIIVANEI 299
Oy      354 ECFEPEG 359
Db      300 KTFEFG 305

RESULT 10
SERB_BACSU STANDARD; PRT: 525 AA.
AC P3136; D32011;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (PGDH).
GN SERA.
DS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID:1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=96349105; PubMed=8760912;
RA Sorokin A.V., Azevedo V., Zumstein E., Galleron N., Ehrlich S.D.,
RA Serror P.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RL the sera and Kds loci cloned in a yeast artificial chromosome."
RM Microbiology 142:2005-2016(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertolo M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borstler S., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fehret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Goldightly E.J., Grandi G.,
RA Gulsepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaser-Blanchard M., Klein C.,
RA Kobayashi Y., Koeter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nockback M.,
RA Noone D., O'Reilly M., Dgawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portwolk S., Prescott A.M.,
RA Presacan E., Pujic P., Punelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadele Y.,
RA Sato T., Scanlon E., Schleich S., Schroeter R., Scofield F.,
RA Sekiuchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Solido B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassartoli A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Weltenegeger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE OF 107-525 FRDM N.A.
RC STRAIN=168 / Marburg;
RX MEDLINE=95020538; PubMed=7934829;
RA Sorokin A.V., Zumstein E., Azevedo V., Ehrlich S.D., Serror P.;

```

```

RT "The organization of the Bacillus subtilis 168 chromosome region
RT between the spoVA and sera genetic loci, based on sequence data.";
RL Mol. Microbiol. 10:385-395(1993).
CC -1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) -> 3-
CC phosphohydroxypyruvate + NADH.
CC -1- ENZYME REGULATION: IN BACTERIA DISPLAYS FEEDBACK INHIBITION BY
CC L-SERINE.
CC -1- PATHWAY: serine biosynthesis; first step.
CC -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L47648; AAC83943.1; -
DR EMBL: 299116; CAB14239.1; -
DR EMBL: L09228; AAA67502.1; -
DR PIR: C69705; C69705.
DR HSP: P33160; ZNAD.
DR Subtilist; BG10509; sera.
DR InterPro: IPR006139; 2-Hacid_DH.
DR InterPro: IPR006140; 2-Hacid_DH.C.
DR InterPro: IPR002912; ACT.
DR InterPro: IPR006236; PGDH.
DR Pfam: PF00368; 2-Hacid_DH.1.
DR Pfam: PF02826; 2-Hacid_DH.C.1.
DR Pfam: PF01842; ACT.1.
DR TIGRFAMS: TIGR01327; PGDH.1.
DR PROSITE: PS00065; D_2_HYDROXYACID_DH.1.1.
DR PROSITE: PS00670; D_2_HYDROXYACID_DH.2.1.
DR PROSITE: PS00671; D_2_HYDROXYACID_DH.3.1.
DR Serine biosynthesis; Oxidoreductase; NAD; Complete proteome.
KW NP_BIND 238 267 NAD (BY SIMILARITY).
FT ACT_SITE 229 229 SUBSTRATE BINDING (BY SIMILARITY).
FT ACT_SITE 258 258 BY SIMILARITY.
FT ACT_SITE 276 276 BY SIMILARITY.
SQ SEQUENCE 525 AA; 57115 MW; B3D0854487DD44FE CRC64;

Query Match 15.88; Score 338; DB 1; Length 525;
Best Local Similarity 31.28; Pred. No. 6.9e-21;
Matches 102; Conservative 54; Mismatches 131; Indels 40; Gaps 9;

Oy 59 GLRPLESNGHTLVTSKDPSPYFERELVDADVISOFPWAVYLPERIAAKNKLA 118
Db 14 GLQPLESDTELVQKNVADAED--ELHTFDLTVRS---ATKVEDLFNKMTSLKIV 66
Oy 119 LTAGIGSDHVDLSAIDRNVTAEVYCNISVAEHVVMILSVRY-----LPSHEW 172
Db 67 GRAGVGDNDIDEDATHGIVINAPNGNTISRAEHFAMISLMRIIPANISVSREM 126
Oy 173 ARKGWNIVADCVSHAY--DLAMHVGTVAAGRIGLAVALRRLAPEDVHLHYDRHLPES 229
Db 127 NRT-----AYVSELYGKTLGIVGLGRIGSEIAORGAFGVTWVHPDPF-LTEE 174
Oy 230 VEKELNLTWATREDYPCVDVYTLNCLPPEFHEHMDTELTLPFRGAVIVTANGKLC 289
Db 175 RAKRIGV-NRTEEVLESADITVHPLRKETKGLNKETTKAGVRLINCARGIT 233
Oy 290 DRDVAVALSGLRAGYAGVWFPAPKDHMPRTMPYNGMTPHISGTLTAAQARYAGT 349
Db 234 DEAALEALELNGHVAAGALDVFEER-PVDNKLVDHVLVATVHTLGASTRKADLVNAAQV 292
Oy 350 RETLECFEGRPT-----RDEY 366
Db 293 SEEVLOFAKGLPYMSAINLPAMTKDEF 319

RESULT 11

```

SERA	METUA	STANDARD:	PRI:	524 AA.
AC	Q58424;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95) (PGDH).			
GN	SERA OR MJ1018			
OS	Methanococcus jannaschii.			
OC	Archaea; Euryarchaeota; Methanococci; Methanococcales;			
OC	Methanocaldococcaceae; Methanocaldococcus.			
OX	NCBI_TaxID=2190;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-JAL-1 / DSM 2661 / ATCC 43067;			
RX	MEDLINE=9633799; PubMed=6688087;			
RA	Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,			
RA	Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,			
RA	Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,			
RA	Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,			
RA	Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,			
RA	Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,			
RA	Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,			
RA	Klenk H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;			
RT	"Complete genome sequence of the methanogenic archaeon, Methanococcus			
RL	jannaschii." Science 273:1058-1073(1996).			
CC	-1- CATALYTIC ACTIVITY: 3-phosphoglycerate + NAD(+) = 3-			
CC	phosphohydroxypruvate + NADH.			
CC	-1- PATHWAY: Serine biosynthesis: first step.			
CC	-1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID			
CC	DEHYDROGENASES FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce#lsb-sib.ch).			
CC	or send an email to license@lsb-sib.ch).			
CC	-----			
DR	EMBL: U67544; AAB99020.1; -			
DR	PIR: A64427; A64427.			
DR	HSSP: P08328; IPSD.			
DR	TIGR: MJ1018; -			
DR	InterPro: IPR006139; 2-Hacid_DH.			
DR	InterPro: IPR006140; 2-Hacid_DH.C.			
DR	InterPro: IPR002912; ACT.			
DR	InterPro: IPR006236; PGDH.			
DR	Pfam: PF00389; 2-Hacid_DH; 1.			
DR	Pfam: PF02826; 2-Hacid_DH.C; 1.			
DR	Pfam: PF01842; ACT; 1.			
DR	TIGRFAMS: TIGR01327; PGDH; 1.			
DR	PROSITE: PS00065; D_2_HYDROXYACID_DH.1; 1.			
DR	PROSITE: PS00670; D_2_HYDROXYACID_DH.2; 1.			
DR	PROSITE: PS00671; D_2_HYDROXYACID_DH.3; 1.			
FT	Serine biosynthesis; Oxidoreductase; NAD; Complete proteome.			
FT	NP_BIND 240 269 NAD (BY SIMILARITY).			
FT	ACT_SITE 231 231 SUBSTRATE BINDING (BY SIMILARITY).			
FT	ACT_SITE 260 260 BY SIMILARITY.			
FT	ACT_SITE 278 278 BY SIMILARITY.			
SO	SEQUENCE 524 AA; 56923 MW; 5C561BB68C10793 CRC64;			
QY	Query Match	15.2%;	Score 323.5;	DB 1; Length 524;
QY	Best Local Similarity	32.1%;	Pred. No. 1.2e-19;	
Db	Matches	96;	Conservative	58; Mismatches 108; Indels 37; Gaps 10;
OY	64	LESGHTLVYTSKDGDPDSYERELV----	DADVVISQFPWPAYLTPERIAKAKNTKLAL	119
OY	11	-----		
OY	18	LEEVGEVEAVT-----GLTKRELEIKIKADAVLVNRS--GTRVTRDVIEKAEKLVIG	68	
OY	120	TAGIGSDHVDQSAIDRNVYTAETVYCNISIAEHVYMMILSVRNY-----LPSHEMA	173	
OY	11	-----		

ID	NAME	FUNCTION	SEQUENCE
Dh	69	RACVGDNDIDVEATKTKGLIVNAPDASSISVAELTMGLMLAARNIPQAPASLKR	EMD 128
Qy	174	RKGGNMADCVSHAVDLEAMHWGVAAGRTGLAVLRRLAPDVHLHTYDRHRLPESVEKE	2333
Dh	129	RRKFRKQ-----ELYKTKTGLVIGLIGGOQVVKRAKAGMNIIGDPY-IPREVAES	179
Qy	234	LNLTWIATREDMYPCV---DVTLMQPLPHEFEHINDETLLKLPFRGAVIYNTRAGL	CD 2909
Dh	180	MGV--ELVDDINELCKRADPFTTLHPLTPKTRHITGREGQALMKNAIYVNCARGGL	LD 236
Qy	291	RDAAVARALESGRIAGDWFPEQAPAKDPKRPRTMPYNGMTPHISGTLTQAARVYAGT	349
Dh	237	EKLVEALKEKGRIRAAALDVFEEEP-PRKNPLLTLDVNIIGTRPGCAS--TEEAQNA	GT 292
RESULT 12			
ID	DHGY_HYPM	STANDARD:	PRT: 321 AA.
AC	P36234:	01-JUN-1994 (Rel. 29, Created)	
DT	01-JUN-1994 (Rel. 29, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Glycylate dehydrogenase (EC 1.1.1.29) (NADH-dependent hydroxypruvate		
DE	reductase) (HPR) (GDH) (Hydroxypruvate dehydrogenase) (Glyoxylate		
DE	reductase).		
OS	Hyphomicrobium methylotrophum.		
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;		
OC	Hyphomicrobiaceae; Hyphomicrobium.		
OX	NCBI_Taxid=84;		
RN	[1]		
RP	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).		
RX	MEDLINE=94160078; Pubmed=8120891;		
RA	Goldberg J.D., Yoshida T., Brick P.;		
RT	"Crystal structure of a NAD-dependent D-glycerate dehydrogenase at		
RT	2.4-A resolution."		
RL	J. Mol. Biol. 236:1123-1140(1994).		
RN	[2]		
RP	CHARACTERIZATION.		
RC	STRAIN=KM146 / GM2;		
RX	MEDLINE=90306024; Pubmed=2114287;		
RA	Izumiy Y., Yoshida T., Kanzaki H., Toki S.-I., Miyazaki S.S.;		
RA	Yamada H.;		
RT	"Purification and characterization of hydroxypruvate reductase from		
RT	a serine-producing methylotroph, Hyphomicrobium methylotroph GM2."		
RL	Eur. J. Biochem. 190:279-284(1990).		
CC	1- FUNCTION: ACTIVE ON HYDROXYPRUVATE AND GLYOXYLATE. OPTIMUM PH IS		
CC	6.8 AND TEMPERATURE IS 45 DEGREES CELSIUS.		
CC	1- CATALYTIC ACTIVITY: (R)-glycerate + NAD(+) = hydroxypruvate +		
CC	NADH.		
CC	1- PATHWAY: Serine pathway.		
CC	1- SUBUNIT: Homodimer.		
CC	1- INDUCTION: By methanol.		
CC	1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID		
CC	DEHYDROGENASES FAMILY.		
CC	PDB: 1GDH; 3I-JAN-94.		
DR	InterPro: IPR006139; 2-Hacid_DH.		
DR	InterPro: IPR006140; 2-Hacid_DH_C.		
DR	Pfam: PF00389; 2-Hacid_DH; 1.		
DR	Pfam: PF02826; 2-Hacid_DH_C; 1.		
DR	PROSITE: PS00065; D_2-HYDROXYACID_DH_1; 1.		
DR	PROSITE: PS00670; D_2-HYDROXYACID_DH_2; 1.		
DR	PROSITE: PS00671; D_2-HYDROXYACID_DH_3; 1.		
KW	Oxidoreductase; NAD; 3D-structure.		
FT	INT MET	0	
FT	DOMAIN	1	99
FT	DOMAIN	100	290
FT	DOMAIN	291	321
FT	NP_BIND	149	177
FT	ACT_SITE	240	240
FT	ACT_SITE	269	269
FT	ACT_SITE	287	287
FT	STRAND	4	7
FT	HELIX	13	20


```

FT TURN 21 22
FT STRAND 24 27
FT TURN 31 32
FT HELIX 36 43
FT TURN 44 45
FT STRAND 48 52
FT TURN 53 54
FT STRAND 57 57
FT HELIX 59 64
FT TURN 67 68
FT STRAND 71 75
FT TURN 80 81
FT STRAND 82 82
FT HELIX 84 89
FT TURN 90 91
FT STRAND 93 95
FT HELIX 102 117
FT TURN 118 119
FT HELIX 120 128
FT TURN 129 130
FT TURN 137 140
FT STRAND 142 142
FT TURN 146 147
FT STRAND 149 153
FT HELIX 157 167
FT TURN 168 170
FT STRAND 172 176
FT HELIX 183 189
FT TURN 190 190
FT STRAND 192 193
FT HELIX 197 203
FT STRAND 206 209
FT TURN 215 219
FT STRAND 221 221
FT HELIX 223 226
FT TURN 227 228
FT TURN 231 232
FT STRAND 233 237
FT HELIX 241 243
FT STRAND 244 244
FT HELIX 246 255
FT TURN 256 256
FT STRAND 263 263
FT TURN 267 270
FT TURN 276 275
FT HELIX 276 278
FT TURN 280 281
FT STRAND 282 284
FT TURN 289 290
FT STRAND 292 292
FT HELIX 293 311
FT TURN 312 313
FT TURN 317 318
FT STRAND 319 319
SQ SEQUENCE 321 AA: EEEF880491A7307 CRC64:

```

Query Match 14.7%; Score 314; DB 1; Length 321;
 Best Local Similarity 30.7%; Pred. No. 3.8e-19;
 Matches 78; Conservative 44; Mismatches 118; Indels 14; Gaps 6;

```

QY 113 KNLKLTAGIGSHVDLSIDRNVTVAEYTCNSISVAEHVMMILSLVR-----NY 166
DB 67 ENIKICISTYSGFDHDLDCCKARGIKVGNAPGVTAETIAIMLLLSARRAGESEKM 126
QY 167 LPSHEMARGKGMNADCVSAHYDLEAMHVGTVAGRIGLAVRLAPEDVHLHYTDHRL 226
DB 127 IRTSMSP---GWELELVGKLDNKTLL--GIYGFSGTGOALAKRAGGDMDIDYDTHRA 181
QY 227 PESVEKELNLTWATREDMTVPCDVVTLNCPLEPTEHMIINDETLKLFKRGAYIVNTARG 286
DB 182 SSSDEASYQATFHDLSLSLSVSGFSLNAPSTPEYTFYFKATIKSLPGCALIVNTARG 241
QY 287 KLCRDVAVARALESGRLAGTAG-DVWFPPQAPKDPWRTMPEYNGMTIHISGTTLLAQAARY 345

```

```

DB 242 DLVNEELVAVALEAGRLA-YAGFDVFAEPNINE-GYYDLPNTLPFHIGSAATQAREDM 299
QY 346 AAGTRELIECFEEG 359
DB 300 AHOANDLIDALEGG 313

```

RESULT 13

```

CTBP_XENLA
ID CTBP_XENLA STANDARD: PRT: 440 AA.
AC Q9YHU0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C-terminal binding protein (CTBP).
GN CTBP.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99078014; PubMed-9858600;
RA Sewalt R.G.A.B., Gunster M.J., van der Vliet J., Salijn D.P.E.,
RA Ote A.P.;
RT "C-terminal binding protein is a transcriptional repressor that
RT interacts with a specific class of vertebrate polycarb proteins."
RL Mol. Cell. Biol. 19:777-787(1999).
CC -!- FUNCTION: CO-REPRESSOR TARGETING DIVERSE TRANSCRIPTION REGULATORS
CC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear (potential).
CC -!- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC DEHYDROGENASES FAMILY. BUT HIGHLY DIVERGENT.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF091554; AAD14596.1; -
CC DR HSSP: P17584; IDXY.
CC DR InterPro: IPR006139; 2-Hacid_DH.
CC DR InterPro: IPR006140; 2-Hacid_DH.C.
CC DR Pfam: PF00389; 2-Hacid_DH.1.
CC DR Pfam: PF02826; 2-Hacid_DH.C.1.
CC DR PROSITE: PS00065; D_2_HYDROXYACID_DH_1; 1.
CC DR PROSITE: PS00670; D_2_HYDROXYACID_DH_2; FALSE NEG.
CC DR PROSITE: PS00671; D_2_HYDROXYACID_DH_3; FALSE NEG.
CC KW Nuclear protein.
SQ SEQUENCE 440 AA: 4776 MW; E1AB044901A00116 CRC64:

```

Query Match 14.5%; Score 309.5; DB 1; Length 440;
 Best Local Similarity 30.5%; Pred. No. 1.4e-18;
 Matches 93; Conservative 52; Mismatches 129; Indels 31; Gaps 6;

```

QY 58 LGLRPYLESGN--HTLVVTSKDGPDVSVE-----RELVDADV 94
DB 13 LGIRPPI-MNGPMHPRRLVALLDGRDCTVEPILKDVATVAFCDASTOEIHEKVLNEAV 71
QY 95 ISGFPMYAVLTPRIIAKMKLTLTGIGSDHVDLSIDRNVTVAEYTCNSISVAEH 154
DB 72 GALTWHITTLTREDLEFKALRIIVRIGSGDNDIDISADDLGIAVCNVPAAVEETADS 131
QY 155 VVMMLSLVENVYLPESHMARKG-----NMNADCVSAHYDLEAMHVGTVAGRIGLAVLR 209
DB 132 TMCIIILNLYKRTTLMHALNEGTVQSVQDREVASGAANKIRGTTGLIIGLRGQVAL 191
QY 210 RLAPFDVHLHYTDHRLPESVEKELNLTWATREDMTVPCDVVTLNCPLEPTEHMIINDE 269

```



```

Db      192 RAKTFGNVFEYDPR-LSDGIERALGLQVRSTLDDLFHSCVYTLHGLNHNHLLNDF 250
OY      270 TLKLFKAGAYIVNTARCKLDRDVAVARALSSGLAGYAGDYWFPQ-APKDHPRTPMYN 328
Db      251 TIKOMRGAFVLTNARGLVDERKALQALKEGRIRGALDYHESEPFSTGPKDAPNL 310
OY      329 GMTPH 333
Db      311 ICTPH 315

RESULT 14
CTPL_HUMAN
ID      CTPL_HUMAN          STANDARD;          PRT;          440 AA.
AC      Q13363;
DT      15-JUL-1998 (Rel. 36, Created)
DT      15-JUL-1999 (Rel. 38, Last sequence update)
DT      15-SEP-2003 (Rel. 42, Last annotation update)
DE      C-terminal binding protein 1 (CTBP1).
GN      CTBP1 OR CTBP.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A., AND SEQUENCE OF 98-108.
RC      TISSUE-B-cell, and Cervical carcinoma;
RX      MEDLINE=9606836; PubMed=7479821;
RA      Schaeper U., Boyd J.M., Verma S., Uhlmann E., Subramanian T.,
RA      Chinnadurai G.;
RT      "Molecular cloning and characterization of a cellular phosphoprotein
RT      that interacts with a conserved C-terminal domain of adenovirus E1A
RT      involved in negative modulation of oncogenic transformation.";
RL      Proc. Natl. Acad. Sci. U.S.A. 92:10467-10471(1995).
RN      [2]
RP      SEQUENCE FROM N.A., AND REVISIONS.
RX      MEDLINE=99078014; PubMed=9858600;
RA      Sewalt R.G.A.B., Gunster M.J., van der Vlag J., Satijn D.P.E.,
RA      Otte A.P.;
RT      "C-terminal binding protein is a transcriptional repressor that
RT      interacts with a specific class of vertebrate polycomb proteins.";
RL      Mol. Cell. Biol. 19:777-787(1999).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE-Brain;
RX      MEDLINE=22386257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muljani S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Wozley K.C., Hale S., Garcia A.M., Gay L.J., Huijyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzyzanski M.I., Skalska U., Smalhus D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [4]
RP      CHARACTERIZATION.
RX      MEDLINE=93178421; PubMed=8440238;
RA      Boyd J.M., Subramanian T., Schaeper U., Ia Regina M., Bayley S.,
RA      Chinnadurai G.;
RT      "A region in the C-terminus of adenovirus 2/5 E1a protein is required

```

```

RT      for association with a cellular phosphoprotein and important for the
RT      negative modulation of T24-ras mediated transformation, tumorigenesis
RT      and metastasis.";
RL      EMBO J. 12:469-478(1993).
CC      -1- FUNCTION: CO-REPRESSOR TARGETING DIVERSE TRANSCRIPTION REGULATORS.
CC      -1- SUBUNIT: INTERACTS WITH THE C-TERMINUS OF ADENOVIRUS E1A PROTEIN,
CC      E1K3 AND CTIP VIA THEIR CONSENSUS MOTIF P-X-(DNS)-L-(STVA).
CC      CAN FORM HOMODIMERS OR HETERODIMERS OF CTBP1 AND CTBP2.
CC      -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC      -1- PTM: PHOSPHORYLATED. THE LEVEL OF PHOSPHORYLATION APPEARS TO BE
CC      REGULATED DURING THE CELL CYCLE.
CC      -1- SIMILARITY: BELONGS TO THE D-ISOMER SPECIFIC 2-HYDROXYACID
CC      DEHYDROGENASES FAMILY. BUT HIGHLY DIVERGENT.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U37408; AAC62822.1; -
DR      EMBL; AF091555; AAD14597.1; -
DR      EMBL; BC011655; AAH11655.1; -
DR      PDB; 1MX3; 18-DEC-02.
DR      GeneW; HGNC:2494; CTBP1.
DR      MIM; 602618; -
DR      GO; GO:0008022; F:protein C-terminus binding activity; TAS.
DR      GO; GO:0008181; F:tumor suppressor; TAS.
DR      GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR      GO; GO:0007048; P:oncogenesis; TAS.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
DR      GO; GO:0008166; P:viral replication; TAS.
DR      InterPro; IPR006139; 2-Hacid_DH.
DR      InterPro; IPR006140; 2-Hacid_DH.C.
DR      Pfam; PF00389; 2-Hacid_DH.1.
DR      Pfam; PF02826; 2-Hacid_DH.C.1.
DR      PROSITE; PS00065; D_2_HYDROXYACID_DH.1; 1.
DR      PROSITE; PS00670; D_2_HYDROXYACID_DH.2; FALSE.NEG.
DR      PROSITE; PS00671; D_2_HYDROXYACID_DH.3; FALSE.NEG.
DR      KEGG; K01101; Phosphorylation; 3D-structure.
DR      KEGG; K01101; Phosphorylation; 3D-structure.
SQ      SEQUENCE 440 AA; 47535 MW; F071DD308385603F CRC64;

Query Match      14.4%; Score 307.5; DB 1; Length 440;
Best Local Similarity 30.8%; Pred. No. 2.1e-18;
Matches 94; Conservative 51; Mismatches 129; Indels 31; gaps 6;

OY      58 LGLRPVLESNG-HTLVVTSKDPDSVFE-----RELVDADVV 94
Db      13 LGVVRPPI-MNCPRLPRPVALLDGRDCTVEPILKDAVTAFCDAQSTQETHEKYLEAV 71
OY      95 ISQFPMYAVLTPERIAKAKNLKLTAGIGSDHVDLQSAIDRNVTVAEVTYCNISVAEH 154
Db      72 GALKVHTITLREDLEKFKALRIIVRGSGFDNDIKSAGDLGIAVCVPAASVEERADS 131
OY      155 VVMILSLVRYTPDSHEWARKG-----WN1ADCVSNAYDEAMNAGVTVAAGRIGLAVLR 209
Db      132 TLCHILMLRYRATWLHQLRSGTQVQSEQIREVASARIRGRTGLTIGIGRGQAVALL 191
OY      210 RLAFEDVHLNTHDRRLPESVEKEJLNTWNTAFRDWYRVCVVVNLCPILPETHEMINDE 269
Db      192 RAKTFGNVFEYDPR-LSDGIERALGLQVRSTLDDLFHSCVYTLHGLNHNHLLNDF 250
OY      270 TLKLFKAGAYIVNTARCKLDRDVAVARALSSGLAGYAGDYWFPQ-APKDHPRTPMYN 328
Db      251 TIKOMRGAFVLTNARGLVDERKALQALKEGRIRGALDYHESEPFSTGPKDAPNL 310
OY      329 GMTPH 333
Db      311 ICTPH 315

```

